

F:228-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: protein 3A #status predicted <P3A>
F:1499-1521/Product: protein 3B #status predicted <P3B>
F:1522-1741/Product: protein 3C #status predicted <P3C>
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 100.0%; Score 129; DB 1; Length 2230;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQRIKEQ 25
DB 960 WLNPKKINLADRMGLSGVQRIKEQ 984

RESULT 6
B82056
Glutathione-regulated potassium-efflux system protein KefB VC2606 [imported] - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82056
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardoon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82056
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-656 <HEI>
A:Cross-references: GB:AE003852; GB:AE004327; PIDN:AAF95747.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2606
A:Map position: 1
C:Superfamily: glutathione-regulated potassium efflux system protein kefC

Query Match 39.9%; Score 51.5; DB 2; Length 656;
Best Local Similarity 54.5%; Pred. No. 13;
Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 2 LNPXXI-NLADRMGLSGVQEI 22
DB 133 LNPXXI-NLADRMGLSGVQEI 154

RESULT 7
S33416
heat shock protein hsp20 - nematode (Nippostrongylus brasiliensis)
C:Species: Nippostrongylus brasiliensis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C:Accession: S33416
R:Twissie, S.; Grigg, M.E.; Ingram, L.; Selkirk, M.E.
submitted to the EMBL Data Library, April 1993
A:Description: The expression of a small heat shock homologue is developmentally regulated
A:Reference number: S33416
A:Accession: S33416
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-172 <TWE>
A:Cross-references: EMBL:X71663; NID:g297865; PIDN:CAA50655.1; PID:g297866
C:Superfamily: alpha-crystallin

Query Match 38.8%; Score 50; DB 2; Length 172;
Best Local Similarity 43.5%; Pred. No. 5.2;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 NPKKINLADRMGLSGVQRIKEQ 25
DB 77 NELKVLDDRDITVEGMQVKT 99

RESULT 8
RDV2AS
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - African swine fever virus
N:Alternate names: ribonucleotide reductase small chain
C:Species: African swine fever virus, ASFV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Sep-2003
C:Accession: B40568
R:Bournnell, M.; Shaw, K.; Yanez, R.J.; Vinuela, E.; Dixon, L.
Virology 184, 411-416, 1991
A:Title: The sequences of the ribonucleotide reductase genes from African swine fever virus
A:Reference number: A40568; MUID:91335775; PMID:1871976
A:Accession: B40568
A:Molecule type: DNA
A:Residues: 1-327 <BOU>
A:Cross-references: GB:M64728

Query Match 38.0%; Score 49; DB 1; Length 327;
Best Local Similarity 31.8%; Pred. No. 14;
Matches 7; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEI 22
DB 145 WMDPARNSLGERLVGPAVEGI 166

RESULT 9
S64321
hypothetical protein YGR030c - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G4068
C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
C:Accession: S64321
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64321
A:Molecule type: DNA
A:Residues: 1-158 <RIE>
A:Cross-references: EMBL:Z72815; NID:g1323007; PID:e243932; PID:g1323008; GSPDB:GN00007
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:POF6; MIPS:YGR030C
A:Cross-references: SGD:S0003262
A:Map position: 7R

Query Match 37.2%; Score 48; DB 2; Length 158;
Best Local Similarity 66.7%; Pred. No. 9.4;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 KKNLADRMGLSGV 19
DB 64 KQINNADRSIGLQGV 78

RESULT 10
S27608
cysteine proteinase tpr - Porphyromonas gingivalis
C:Species: Porphyromonas gingivalis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 04-Mar-2000
C:Accession: S27608
R:Bourgau, G.; Lapointe, H.; Ploquin, P.; Mayrand, D.
submitted to the EMBL Data Library, February 1992

A;Description: Cloning, expression and sequencing of a protease gene (tpr) from Porphyromonas gingivalis

A;Reference number: S27608

A;Accession: S27608

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-482 <BOU>

A;Cross-references: EMBL:M84471; NID:g150851; PIDN:AA25652.1; PID:g150852

C;Superfamily: Porphyromonas gingivalis cysteine proteinase tpr

Query Match 36.8%; Score 47.5; DB 2; Length 482;

Best Local Similarity 33.3%; Pred. No. 36;

Matches 8; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

QY 1 LNPCKINLADRMGLSGVQEIKE 24
||| : : : ||| : : : :
80 WSNPRTTDIVERVLG-SSMQDLK 102

Db

RESULT 11

G82971

Probable ferredoxin PA5399 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C;Accession: G82971

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Baderman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen

A;Reference number: AB2950; MUID:20437337; PMID:10984043

A;Accession: G82971

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-653 <STO>

A;Cross-references: GB:AE004952; GB:AE004091; NID:g9951718; PIDN:AAG08784.1; GSPDB:GN001

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA5399

Query Match 36.8%; Score 47.5; DB 2; Length 653;

Best Local Similarity 45.5%; Pred. No. 50;

Matches 10; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 2 LNPCKINLADRMGLSGVQEIKE 23
||| : : : ||| : : : :
264 LNPCKL-IDQWVIGLAGNDK 284

Db

RESULT 12

B25199

Heat shock 16K protein 2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 13-Aug-1999

C;Accession: B25199

R;Jones, D.; Ruesnak, R.H.; Kay, R.J.; Candido, E.P.M.

J. Biol. Chem. 261, 12006-12015, 1986

A;Title: Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis elegans

A;Reference number: A92555; MUID:86304344; PMID:3017958

A;Accession: B25199

A;Molecule type: DNA

A;Residues: 1-145 <ON>

A;Cross-references: GB:M14334; NID:g156338; PIDN:AAA28071.1; PID:g156340

C;Superfamily: alpha-crystallin

Query Match 36.4%; Score 47; DB 2; Length 145;

Best Local Similarity 55.6%; Pred. No. 12;

Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 KINLADRMGLSGVQEIKE 23
||| : : : ||| : : : :
65 KINLGDRTLSIQGEQLK 82

Db

A;Molecule type: DNA
A;Residues: 1-342 <PAR>
A;Cross-references: GB:AL139077; GB:AL111168; NID:G6968444; PIDN:CAB73388.1; PID:G696856
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: waaC; Cj1133
C;Superfamily: ADP-heptose-LPS heptosyltransferase II
Query Match 35.7%; Score 46; DB 2; Length 342;
Best Local Similarity 42.9%; Pred. NO. 42;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 1 WLNPKKINLADRMGLSGVQE 21
DB 220 WGNVXEYEFKEVLNLSGIDE 240

Search completed: March 15, 2004, 14:07:01
Job time : 12.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:26:12 ; Search time 7.05882 Seconds
(without alignments)
184.415 Million cell updates/sec

Title: US-09-171-432a-48
Perfect score: 129
Sequence: 1 WLNPKINLADRLGLSGVQEIKEQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	2226	1 POLG_HPAV2	P26580 hepatitis a
2	129	100.0	2226	1 POLG_HPAV4	P26581 hepatitis a
3	129	100.0	2226	1 POLG_HPAV8	P26582 hepatitis a
4	129	100.0	2227	1 POLG_HPAVH	P08617 hepatitis a
5	129	100.0	2227	1 POLG_HPAVJ	P06441 hepatitis a
6	129	100.0	2227	1 POLG_HPAVM	P13901 hepatitis a
7	129	100.0	2230	1 POLG_HPAVS	P14553 simian hepa
8	50	38.8	172	1 H820_NIEBR	Q07160 nipprostrong
9	49	38.0	327	1 RIR2_ASFM2	P26713 african swi
10	48	37.2	158	1 POF6_YEAST	P53218 saccharomyc
11	48	37.2	334	1 RIR2_ASFB7	P42492 african swi
12	48	37.2	622	1 LEPA_STRCO	Q9rdc9 streptomyc
13	47.5	36.8	481	1 TRR_FORGI	P25806 porphyromon
14	47.5	36.8	603	1 LEPA_SYNEL	Q8dm40 synechococc
15	47	36.4	145	1 H812_CABEL	P06582 caenorhabdi
16	47	36.4	145	1 RIR2_CABEL	P42170 caenorhabdi
17	46.5	36.0	146	1 R113_BORBU	O51314 borrelia bu
18	46	35.7	518	1 Y448_MYCPN	P75066 mycoplasma
19	46	35.7	632	1 Y449_MYCPN	P75065 mycoplasma
20	45.5	35.3	201	1 H151_CAMJE	Q9pmv4 campylobact
21	45	34.9	622	1 LEPA_STRAW	Q82bz3 streptomyc
22	45	34.9	1042	1 SVI_BORBU	O51773 borrelia bu
23	44	34.1	152	1 ENR9_BPT3	P20314 bacterioph
24	44	34.1	165	1 YC36_FORPU	P51273 porphyra pu
25	44	34.1	639	1 DNAK_SHEON	Q8eht7 shewanella
26	44	34.1	646	1 LEPA_MYCLE	P53530 mycobacteri
27	44	34.1	653	1 LEPA_MYCTU	P71739 mycobacteri
28	44	34.1	985	1 ENV_SVU1	P32073 simian foam
29	44	34.1	1770	1 R115_YEAST	P43565 saccharomyc
30	43.5	33.7	285	1 RFP32_BUCAP	Q8ka76 buchnera ap
31	43.5	33.7	388	1 ALR_MYCLE	P38056 mycobacteri
32	43.5	33.7	441	1 SDHD_FUSNN	Q8rfx6 fusobacteri
33	43.5	33.7	482	1 HOXA_ALCEU	P29267 alcaligenes

ALIGNMENTS

RESULT 1

POLG_HPAV2

ID POLG_HPAV2 STANDARD; PRT; 2226 AA.

AC P26580;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase

DE P3D (EC 2.7.7.48)]

OS Hepatitis A virus (strain 24a).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12094;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91162758; PubMed=1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,

RA Cromeans T., Jansen P.W.;

RT Antigenic and Genetic Variation in cytopathic hepatitis A virus

RT Variants arising during persistent infection: evidence for genetic

RT recombination.;

RL J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC (RNA) (N).

CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M59810; AAA45468.1; -

CC MEROPS; C03.005; -

DR InterPro; IPR004004; Calici_pol_hel.

DR InterPro; IPR009003; Cys_Ser_trypsin.

DR InterPro; IPR000605; RNA_helicase.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR001205; RNA_pol_P3D.

DR InterPro; IPR007094; RNA_pol_Psvir.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA_helicase; 1.

DR PRINTS; PR00918; CALICIVIRUSN.

DR PolyProtein; Coat protein; Core protein; Transferase;

DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.

CHAIN 1 23 COAT PROTEIN VP4 (P1A).

FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).

FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).

FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 COAT PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D684E2BF CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
 DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980
 RESULT 2
 POLG_HPAV4
 ID POLG_HPAV4 STANDARD; PRT; 2226 AA.
 AC P26581;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE P3D (EC 2.7.7.48).
 OS Hepatitis A virus (strain 43c).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M59809; AAA45469.1; --
 CC MEROPS: C03.005; --
 CC InterPro: IPR004004; Calici_pol_hel.
 CC InterPro: IPR009003; Cys_Ser_trypsin.
 CC InterPro: IPR000605; RNA_helicase.
 CC InterPro: IPR007095; RNA_pol_DS_PS.
 CC InterPro: IPR001205; RNA_pol_P3D.
 CC InterPro: IPR007094; RNA_pol_PSVir.
 CC InterPro: IPR008975; Viral_cap_coat.
 CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
 CC InterPro: IPR00910; RNA_helicase; 1.
 CC PRINTS: PR00918; CALICIVIRUS.
 CC Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 CC COAT PROTEIN VP4 (P1A).
 FT CHAIN 1 23

FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9BF75 CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
 DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980
 RESULT 3
 POLG_HPAV8
 ID POLG_HPAV8 STANDARD; PRT; 2226 AA.
 AC P26582;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE P3D (EC 2.7.7.48).
 OS Hepatitis A virus (strain 18f).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M59808; AAA45467.1; --
 CC MEROPS: C03.005; --
 CC InterPro: IPR004004; Calici_pol_hel.
 CC InterPro: IPR009003; Cys_Ser_trypsin.
 CC InterPro: IPR000605; RNA_helicase.
 CC InterPro: IPR007095; RNA_pol_DS_PS.
 CC InterPro: IPR001205; RNA_pol_P3D.
 CC InterPro: IPR007094; RNA_pol_PSVir.
 CC InterPro: IPR008975; Viral_cap_coat.
 CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
 CC InterPro: IPR00910; RNA_helicase; 1.
 CC PRINTS: PR00918; CALICIVIRUS.

KW Polyprotein; Coat protein; Core protein; Transferase;
FT RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;
Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980
RESULT 4
ID POLG HPAHV STANDARD; PRT; 2227 AA.
AC P08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031696;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14114; AAA45475.1; -;
DR EMBL; M14707; AAA45465.1; -;
DR EMBL; M14707; AAA45466.1; ALT_INIT.
DR EMBL; M16632; AAA45471.1; -;
DR PIR; A03905; A03905.
DR PIR; A25981; GNNYHM.
DR PIR; A94149; GNNYMK.
DR PDB; 1HAV; 23-DEC-96.
DR MEROPS; C03.005; -;
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR006005; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR01205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol_1.
DR Pfam; PF00910; RNA_helicase_1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT CHAIN 77 77
FT CHAIN 764 764
FT CHAIN 821 821
FT CHAIN 1052 1052
FT CHAIN 1062 1062
FT CHAIN 1118 1118
FT CHAIN 1151 1151
FT CHAIN 1163 1163
FT CHAIN 1277 1277
FT CHAIN 1500 1500
FT CHAIN 1805 1805
FT CHAIN 1930 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;
Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980
RESULT 5
POLG HPAVL STANDARD; PRT; 2227 AA.
ID POLG HPAVL
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain LA).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatocivirus.
 OX NCBI_TaxID=12099;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85190549; PubMed=2986127;
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
 RA Marryweather J., van Nest G., Dina D.,
 RT "Primary structure and gene organization of human hepatitis A virus."
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; K02990; AAA45472.1; -.
 CC PIR; A03903; GNNYHR.
 CC MEROPS; C03.005; -.
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUS.
 CC Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1076
 FT CHAIN 1077 1422
 FT CHAIN 1423 1484
 FT CHAIN 1485 1507
 FT CHAIN 1508 1678
 FT CHAIN 1679 2227
 SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
 Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 6

POLG HPVAM

ID POLG HPVAM STANDARD; PRT; 2227 AA.

AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;

AC Q81090; Q81091; Q81092; Q81093;

DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain MBB).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatocivirus.
 OX NCBI_TaxID=12100;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
 RA Deinhardt F.,
 RT "The entire nucleotide sequence of the genome of human hepatitis A
 RT virus (isolate MBB)."
 RL Virus Res. 8:153-171(1987).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
 CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M20273; AAA45474.1; -.
 CC MEROPS; C03.005; -.
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUS.
 CC Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 7

POLG HPAVS STANDARD; PRT; 2230 AA.
ID P14553;
AC 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683 (1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428 (1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D00924; BAA00766.1; -;
CC EMBL; X15461; CRA33490.1; -;
CC PIR; A30470; GNNYSA.
CC
CC InterPro; IPR004004; Calici pol hel.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR000805; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_Psvir.
CC InterPro; IPR008975; Viral cap coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUS.
CC PolyProtein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
FT CHAIN 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
SQ

Query Match 100.0%; Score 129; DB 1; Length 2230;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKINLADRMILGLSGVQEIKEQ 25
DB 960 WLNPKINLADRMILGLSGVQEIKEQ 984

RESULT 8
HS20 NIPBR STANDARD; PRT; 172 AA.
AC Q07160;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Heat shock protein homolog (HSP20).
GN HSP20.
OS Nippostrongylus brasiliensis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Heligmonellidae; Nippostrongylinae;
OC Nippostrongylus.
OX NCBI_TaxID=27835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94081860; PubMed=8259127;
RA Tweedie S., Grigg M.E., Ingram L., Selkirk M.E.;
RT "The expression of a small heat shock protein homologue is
RT developmentally regulated in Nippostrongylus brasiliensis.";
RL Mol. Biochem. Parasitol. 61:149-154 (1993).
CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X71663; CAA50655.1; -;
CC PIR; S33416; S33416.
CC InterPro; IPR001436; Crystallin_alpha.
CC InterPro; IPR002088; Hsp20.
CC InterPro; IPR008978; HSP20_chap.
CC Pfam; PF00011; HSP20; 1.
CC PRINTS; PR00299; ACRYSTALLIN.
CC PROSITE; PS01031; HSP20; 1.
KW Heat shock.
SQ SEQUENCE 172 AA; 20227 MW; 2CDAA711CB60B1C0 CRC64;

Query Match 38.8%; Score 50; DB 1; Length 172;
Best Local Similarity 43.5%; Pred. No. 2;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 NPXKINLADRMILGLSGVQEIKEQ 25
DB 77 NELKVLDDRLTVEGMQSVKTE 99

RESULT 9
RIR2 ASFM2 STANDARD; PRT; 327 AA.
ID RIR2 ASFM2
AC P267I3;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase).
OS African swine fever virus (isolate Malawi Lil 20/1) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.

```

OX NCBI_TaxID=10500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91335775; PubMed=1871976;
RA Boursnell M., Shaw K., Yanez R.J., Vinuela E., Dixon L.;
RT "The sequences of the ribonucleoside diphosphate genes from African
RT swine fever virus show considerable homology with those of the
RT orthopoxvirus, vaccinia virus.";
RL Virology 184:411-416(1991).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thiorodexin + H(2)O = ribonucleoside diphosphate + reduced
CC thiorodexin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large and a small chain.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M64728; -, NOT ANNOTATED_CDS.
CC PIR; B40568; RDVZAS.
CC HSPSP; P11157; IXSM.
CC InterPro; IPR000358; Ribonucl_redcase.
CC Pfam; PF00268; ribonuc_red_sm; 1.
CC PROSITE; PS00368; RIBORED_SMALL; 1.
CC Oxidoreductase; DNA replication; Metal-binding; Iron.
KW FT METAL 70 70 IRON 1 (BY SIMILARITY).
FT METAL 101 101 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 104 104 IRON 1 (BY SIMILARITY).
FT METAL 164 164 IRON 2 (BY SIMILARITY).
FT METAL 198 198 IRON 2 (BY SIMILARITY).
FT METAL 201 201 IRON 2 (BY SIMILARITY).
FT ACT SITE 108 108 BY SIMILARITY.
SQ SEQUENCE 327 AA; 38966 MW; E78508DB1978F4B0 CRC64;

Query Match 38.0%; Score 49; DB 1; Length 327;
Best Local Similarity 31.8%; Pred. No. 5.7;
Matches 7; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEI 22
   : : : : : : : : : : : :
Db 145 WMDPARNISGERLVGFAAVEGI 166

RESULT 10
POP6 YEAST
ID POP6 YEAST STANDARD; PRT; 158 AA.
AC P53218;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribonucleases P/MRP protein subunit POP6 (EC 3.1.26.5) (RNases P/MRP
DE 18.2 kDa subunit) (RNA processing protein POP6).
GN POP6 OR YGR030C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=9743548; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.";

```

```

RL Yeast 13:1077-1090(1997).
CC -!- FUNCTION: Component of ribonuclease P, a protein complex that
CC generates mature tRNA molecules by cleaving their 5' ends. Also a
CC component of RNase MRP.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC extra-nucleotide from tRNA precursor.
CC -!- SUBUNIT: Component of nuclear RNase P and RNase MRP RNase P
CC ribonucleoproteins. RNase P consists of a RNA moiety and at least
CC 8 protein subunits; POP1, POP3, POP4, POP5, POP6, POP7, POP8 and
CC RPPI.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z72815; CA97018.1; -.
CC PIR; S64321; S64321.
CC GerMOnline; I41342; -.
CC SGD; S0003262; POP6.
CC GO; GO:0005655; C:nuclear ribonuclease P complex; IDA.
CC GO; GO:0000172; C:ribonuclease MRP complex; IDA.
CC GO; GO:0000171; F:ribonuclease MRP activity; IDA.
CC GO; GO:0004526; F:ribonuclease P activity; IDA.
KW Hydrolase; Nuclear protein; tRNA processing; Coiled coil.
FT DOMAIN 51 71 COILED COIL (POTENTIAL)
SQ SEQUENCE 158 AA; 18210 MW; 6C27A73FAD521181 CRC64;

Query Match 37.2%; Score 48; DB 1; Length 158;
Best Local Similarity 66.7%; Pred. No. 3.8;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 KKNLADRMGLSGV 19
   : : : : : : : : : :
Db 64 KQINMADRSLGLQV 78

RESULT 11
RIR2 ASF57
ID RIR2 ASF57 STANDARD; PRT; 334 AA.
AC P42492;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase).
GN P334L.
OS African swine fever virus (strain BA71V) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10498;
RN [1]
RP SEQUENCE FROM N.A.
RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,
RA Rodriguez J.F., Vinuela E.;
RT "Analysis of the complete nucleotide sequence of African swine fever
RT virus.";
RL Virology 208:249-278(1995).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thiorodexin + H(2)O = ribonucleoside diphosphate + reduced
CC thiorodexin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large and a small chain.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M84471; AAA25652.1; --
DR EMBL; AE017175; AAQ66170.1; --
DR PIR; S27608; S27608.
DR MEROPS; C01.090; --
DR TIGR; PG1055; --
DR InterPro; IPR001300; Peptidase C2.
DR InterPro; IPR000169; SHprot_acsite.
DR SMART; SM00230; Cyspc; 1.
DR PROSITE; PS0203; CALPAIN CAT; 1.
DR PROSITE; PS00640; THIOL PROTEASE ASN; FALSE_NEG.
DR PROSITE; PS01139; THIOL PROTEASE CYS; 1.
DR PROSITE; PS00639; THIOL PROTEASE HIS; FALSE_NEG.
KW Hydrolase; thiol protease; Virulence; Complete proteome.
FT DOMAIN 169 481 CALPAIN CATALYTIC.
FT ACT_SITE 229 229 BY SIMILARITY.
FT ACT_SITE 406 406 BY SIMILARITY.
FT ACT_SITE 426 426 BY SIMILARITY.
FT ACT_SITE 274 274 A -> AN (IN REP. 1).
FT CONFLICT 274 274
SQ SEQUENCE 481 AA; 54991 MW; 9CC973EF01908383 CRC64;

Query Match 36.8%; Score 47.5; DB 1; Length 481;
Best Local Similarity 33.3%; Pred. No. 15;
Matches 8; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

QY 1 WLNPKKINLADRMGLSGVQEIKE 24
DB 80 WSNPRITDIVERVLG-SSMQDLK 102

RESULT 14

LEPA SYNEL
ID LEPA SYNEL STANDARD; PRT; 603 AA.
AC QSDM20;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GTP-Binding protein lepa.
GN LEPA OR TLR0304
OS Synecchococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecchococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Kato H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC Lepa subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AP005369; BAC07857.1; --
DR HAMAP; MF_00071; -- 1.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG C.
DR InterPro; IPR009022; EFG III V.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR006297; Lepa.

DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR01393; lepa; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACITOR GTP; 1.
KW GTP-binding; Complete proteome.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 82 86 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 603 AA; 67578 MW; ACBBACD1407759AB CRC64;
Query Match 36.8%; Score 47.5; DB 1; Length 603;
Best Local Similarity 52.2%; Pred. No. 19;
Matches 12; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 6 KINLADRMGLSGV---QEIKEQ 25
DB 22 KSTLADRLQLTGTVDPREMEKQ 44

RESULT 15

HS12 CAEL STANDARD; PRT; 145 AA.
ID HS12 CAEL
AC P08582;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heat shock protein HSP16-2.
GN HSP16-2 OR Y46H3A.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86304344; PubMed=3017958;
RA Jones D., Ruesnak R.H., Kay R.J., Cardido E.P.M.;
RT "Structure, expression, and evolution of a heat shock gene locus in
RT Caenorhabditis elegans that is flanked by repetitive elements.";
RL J. Biol. Chem. 261:12006-12015(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Stoneking T., Wohlmann P., Lennox S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M14334; AAA28071.1; --
DR EMBL; AC006774; AAF60615.1; --
DR PIR; B25199; B25199.
DR WormPep; Y46H3A.3; CB22002.
DR InterPro; IPR001436; Crystallin_alpha.
DR InterPro; IPR002668; HSP20.
DR InterPro; IPR008978; HSP20_chap.
DR Pfam; PF00011; HSP20; 1.
DR PRINTS; PR00299; ACRYSTALLIN.

DR PROSITE; PS01031; HSP20; 1.
KW Heat shock; Multigene family.
SQ SEQUENCE 145 AA; 16242 MW; 8A73449F99161889 CRC64;
Query Match 36.4%; Score 47; DB 1; Length 145;
Best Local Similarity 55.6%; Pred. No. 5;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 6 KINLADRMGLSLGVQEI 23
Db ||||| : |||||
65 KINLDGRTLSTQGEQLK 82

Search completed: March 15, 2004, 14:01:02
Job time : 8.05882 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:49:12 ; Search time 36.4706 Seconds
(without alignments)
216.283 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	251	12 Q9ENP8	Q9enp8 hepatitis a
2	129	100.0	251	12 Q9ENN9	Q9enn9 hepatitis a
3	129	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
4	129	100.0	251	12 Q9ENQ2	Q9enq2 hepatitis a
5	129	100.0	251	12 Q9ENN2	Q9enn2 hepatitis a
6	129	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
7	129	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
8	129	100.0	251	12 Q9ENR1	Q9enr1 hepatitis a
9	129	100.0	251	12 Q9ENP4	Q9enp4 hepatitis a
10	129	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
11	129	100.0	251	12 Q9ENNA	Q9enn4 hepatitis a
12	129	100.0	251	12 Q9ENNA	Q9enn4 hepatitis a
13	129	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
14	129	100.0	251	12 Q9ENQ3	Q9enq3 hepatitis a
15	129	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
16	129	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a

17	129	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
18	129	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
19	129	100.0	251	12 Q9ENP6	Q9enp6 hepatitis a
20	129	100.0	251	12 Q9ENQ8	Q9enq8 hepatitis a
21	129	100.0	251	12 Q9ENP3	Q9enp3 hepatitis a
22	129	100.0	251	12 Q9ENQ0	Q9enq0 hepatitis a
23	129	100.0	1124	12 Q84780	Q84780 hepatitis a
24	129	100.0	1161	12 Q05794	Q05794 hepatitis a
25	129	100.0	2216	12 Q9WMA2	Q9wma2 hepatitis a
26	129	100.0	2218	12 Q67824	Q67824 hepatitis a
27	129	100.0	2218	12 Q67817	Q67817 hepatitis a
28	129	100.0	2225	12 Q9DL32	Q9dl32 hepatitis a
29	129	100.0	2227	12 Q9WMA0	Q9wma0 hepatitis a
30	129	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
31	129	100.0	2227	12 Q67825	Q67825 hepatitis a
32	129	100.0	2227	12 Q9WM99	Q9wm99 hepatitis a
33	129	100.0	2227	12 Q8QV03	Q8qv03 hepatitis a
34	129	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
35	129	100.0	2227	12 Q67826	Q67826 hepatitis a
36	129	100.0	2227	12 Q8VON6	Q8von6 hepatitis a
37	129	100.0	2227	12 Q91FH5	Q91fh5 hepatitis a
38	129	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
39	125	96.9	251	12 Q9ENP0	Q9enp0 hepatitis a
40	125	96.9	251	12 Q9ENN5	Q9enn5 hepatitis a
41	125	96.9	251	12 Q9ENN7	Q9enn7 hepatitis a
42	124	96.1	2225	12 Q9DWR1	Q9dwr1 hepatitis a
43	124	96.1	2227	12 Q8QRI6	Q8qri6 hepatitis a
44	121	93.8	251	12 Q9ENR0	Q9enr0 hepatitis a
45	121	93.8	251	12 Q9ENN8	Q9enn8 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENP8 PRELIMINARY; PRT; 251 AA.
ID Q9ENP8
AC Q9ENP8, 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A304;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047665; BAB12173.1; -.
FT NON_TER
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 2

Q9ENN9 PRELIMINARY; PRT; 251 AA.
ID Q9ENN9
AC Q9ENN9, 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

```

RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047661; BAB12169.1; -.
FT NON TER 1
FT NON TER 251
SQ SEQUENCE 251 AA; 28723 MW; 558A2D3664C7343C CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 5
Q9ENN2 PRELIMINARY; PRT; 251 AA.
AC Q9ENN2;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
ON NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A9;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047681; BAB12189.1; -.
FT NON TER 1
FT NON TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 6
Q9ENP2 PRELIMINARY; PRT; 251 AA.
AC Q9ENP2;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
ON NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A5;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON TER 1
FT NON TER 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;

```

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 7

Q9ENQ6 PRELIMINARY; PRT; 251 AA.

ID Q9ENQ6; AC Q9ENQ6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A162;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; --
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6D3B CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 8

Q9ENR1 PRELIMINARY; PRT; 251 AA.

ID Q9ENR1; AC Q9ENR1;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A1;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1; --
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 9

Q9ENP1 PRELIMINARY; PRT; 251 AA.

ID Q9ENP1; AC Q9ENP1;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

ID Q9ENP1 PRELIMINARY; PRT; 251 AA.

AC Q9ENP1; AC Q9ENP1;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A503;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; --
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 10

Q9ENP4 PRELIMINARY; PRT; 251 AA.

ID Q9ENP4; AC Q9ENP4;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A414;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047669; BAB12177.1; --
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 5A315F46C830D214 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 11

Q9ENQ9 PRELIMINARY; PRT; 251 AA.

ID Q9ENQ9; AC Q9ENQ9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

```

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A159;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 12
O9ENN4 PRELIMINARY; PRT; 251 AA.
ID Q9ENN4
AC Q9ENN4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A17;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047679; BAB12187.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 13
O9ENP5 PRELIMINARY; PRT; 251 AA.
ID Q9ENP5
AC Q9ENP5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12176.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 14
Q9ENQ3 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ3
AC Q9ENQ3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A204;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047660; BAB12168.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

RESULT 15
Q9ENP7 PRELIMINARY; PRT; 251 AA.
ID Q9ENP7
AC Q9ENP7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A306;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144
```

Db 120 WLNPKKINLADRLGLSGVQEIKEQ 144

Search completed: March 15, 2004, 14:05:23
Job time : 37.4706 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:24:17 ; Search time 44.2353 Seconds
(without alignments)
127.748 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101
Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	20	AAW42926	Immunogen
2	101	100.0	21	AAB69443	Synthetic
3	101	100.0	854	AAP50287	Hepatitis
4	101	100.0	1077	AAW95559	A partial
5	101	100.0	1091	AAW32426	Translate
6	101	100.0	2227	AAW05697	Attenuate
7	101	100.0	2227	AAW34074	Hepatitis
8	101	100.0	2227	AAW18609	Amino aci
9	101	100.0	2227	AAW18607	Amino aci
10	101	100.0	2227	AAW18608	Amino aci
11	101	100.0	2227	AAE19899	Hepatitis
12	101	100.0	2227	ABG31729	Attenuate
13	101	100.0	2227	ABG31727	Wild-type
14	101	100.0	2227	ABG31728	Hepatitis
15	101	100.0	2227	ABU08640	Attenuat
16	101	100.0	2227	ABU08641	Attenuat
17	101	100.0	2227	ABU08639	Wild type
18	101	100.0	2227	ABW00350	Hepatitis
19	98	97.0	2227	AAP60066	Sequence
20	97	96.0	366	AAP50230	Sequence
21	97	96.0	993	AAP50116	Sequence
22	97	96.0	993	AAP50231	Sequence
23	49.5	49.0	20	AAW76539	Plasmid p
24	47	46.5	515	ABU40548	Protein e
25	47	46.5	518	AAW27340	Group B S

26	47	46.5	816	2	AAW27339	Group B S
27	47	46.5	846	5	ABP30554	Streptoco
28	47	46.5	847	5	ABP29773	Streptoco
29	47	46.5	847	5	ABP27030	Streptoco
30	43	42.6	171	5	ABP40360	Staphyloc
31	43	42.6	259	6	ABM68921	Phototrab
32	43	42.6	720	6	ABU31223	Protein e
33	43	42.6	1108	6	ABU18908	Immunogen
34	42	41.6	20	2	AAW42925	Immunogen
35	42	41.6	21	4	ABM69442	Synthetic
36	42	41.6	125	5	ABM49486	Listeria
37	42	41.6	172	5	ABP25691	Novel hum
38	42	41.6	503	4	ABG15762	Putative
39	41	40.6	144	4	ABM96238	E. faeciu
40	41	40.6	165	7	ADC97403	Protein e
41	41	40.6	173	4	ABM49279	Protein e
42	41	40.6	238	4	ABM68889	Drosophil
43	41	40.6	465	6	ABU24044	Protein e
44	41	40.6	1040	7	ADD45886	Rat Prote
45	41	40.6	2431	2	AAW25138	SFV4 non-

ALIGNMENTS

RESULT 1
AAW42926
ID AAW42926 standard; peptide; 20 AA.
XX
AC AAW42926;
XX
AC AAW42926;
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1319.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX
PI (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
WPI; 1997-535831/49.

Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal.
Claim 18; Page 112; 140pp; English.
Peptides AAW42922-30 are immunogenic peptides corresponding to immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are substantially similar to a portion of the amino acid sequence of the P2A protein of HAV corresponding to amino acids 792-980. The present peptide is derived from amino acids 834-853, and has a reactivity of 27.1% with acute sera. Compositions containing the peptides can also be used to induce an immune response to HAV in a mammal. The peptides can also be used to detect the presence of antibodies against HAV in mammalian serum. The peptides can also be used to make an antibody against HAV by administering the peptide to a mammal
Sequence 20 AA;

Query Match	100.0%;	Score 101;	DB 2;	Length 20;
Best Local Similarity	100.0%;	Pred. No. 1.1e-09;		
Matches	20;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	FSQAKISLFYTEHEIMKFS	20	
DB	1	FSQAKISLFYTEHEIMKFS	20	
RESULT 2				
AAB69443	standard; peptide; 21 AA.			
AC	AAB69443;			
DT	20-APR-2001 (first entry)			
DE	Synthetic HAV P2A peptide, SEQ ID NO: 43.			
KW	Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;			
KW	antigen; major structural capsid polypeptide; HAV antibody detection.			
XX				
OS	Hepatitis A virus.			
OS	Synthetic.			
XX				
XX	WO200105824-A2.			
PN				
XX				
PD	25-JAN-2001.			
XX				
XX	14-JUL-2000; 2000WO-US019267.			
XX				
PR	15-JUL-1999; 99US-0144412P.			
XX				
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.			
PA				
XX				
PI	Fields HA, Khudyakov YE;			
XX				
DR	WFI; 2001-112681/12.			
XX				
PT	Synthetic peptides used as antigen sources for enzyme immunoassays			
PT	detecting anti-hepatitis A virus and as vaccines.			
XX				
PS	Claim 13; Page 95; 130pp; English.			
XX				
CC	The present sequence is one of a number of synthetic peptides which are			
CC	immunoreactive with hepatitis A virus (HAV) antibodies. The peptides			
CC	comprise antigenic epitopes of the major structural capsid polypeptides			
CC	or non-structural polypeptides of HAV with one or more glutamine			
CC	molecules at the carboxy end of the peptide. The peptides are used to			
CC	detect the presence of antibodies against HAV in mammalian serum to			
CC	detect the presence of HAV in a human or animal through the binding of			
CC	the peptide to an antibody, to detect acute phase infection by detecting			
CC	IGM antibodies in mammalian serum and detecting convalescence in a			
CC	mammal. The peptides are used to detect or quantify HAV antibodies in			
CC	samples in clinical or research-based assays using immunoblotting,			
CC	fluorescence in situ hybridisation analysis, gel-mobility shift assays,			
CC	tracking of radioactive or bioluminescent markers, chromatography or			
CC	electrophoresis. The peptides are used to induce an immune response to			
CC	HAV when administered to a human or animal. Glutamine at the carboxy end			
CC	of the peptides enhances the IGM antibody reactivity			
XX				
SQ	Sequence 21 AA;			
Query Match 100.0%; Score 101; DB 4; Length 21;				
Best Local Similarity 100.0%; Pred. No. 1.1e-09;				
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	FSQAKISLFYTEHEIMKFS	20	
DB	1	FSQAKISLFYTEHEIMKFS	20	

XX PF 06-JUN-1995; 95US-00468926.
XX PR 30-SEP-1983; 83US-00537911.
XX PR 27-SEP-1984; 84US-00654942.
XX PR 06-OCT-1988; 88US-00256135.
XX PR 06-NOV-1991; 91US-00788262.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Emerson SU, Purcell RH;
XX PI WPI; 1999-094412/08.
XX DR N-PSDB; AAX01006.
XX CC Chimeric hepatitis A virus strains - with P2 region from cell-culture-
XX PT adapted strain in wild-type genome.
XX PS Disclosure; Fig 7A-L; 36pp; English.
XX CC The present sequence represents a partial hepatitis A virus (HAV)
XX CC protein. The specification describes a DNA construct consisting of a wild
XX CC type HAV genome in which the P2 region is replaced by the P2 region from
XX CC a cell-culture-adapted HAV strain. The construct is used to demonstrate
XX CC that mutations in the P2 region of a cell-culture-adapted HAV strain are
XX CC sufficient for establishment of infection and accelerated growth in cell
XX CC culture
XX CC
XX CC Sequence 1077 AA;
XX CC
Query Match 100.0%; Score 101; DB 2; Length 1077;
Best Local Similarity 100.0%; Pred. No. 8.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSOAKISLFYTEHEIMKFS 20
Db 1057 FSOAKISLFYTEHEIMKFS 1076
RESULT 5
AAR32426
ID AAR32426 standard; protein; 1091 AA.
XX AC AAR32426;
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 17-DEC-2001 (revised)
XX DT 10-JUN-1993 (first entry)
XX CC Translated from 5' region of Hepatitis A Virus genomic clone.
XX CC
XX KW HAV HM-175; chronic liver disease; picornavirus.
XX OS Hepatitis A virus.
XX FH Key Location/Qualifiers
XX FT Region 1. .711
XX FT /note= "X's correspond to nonsense codons, i.e. this
XX FT region is not an ORF"
XX FT 238. .1091
XX FT /label= ORF
XX FT /note= "second putative initiation codon at position 240"
XX PN USN7788262-N.
XX XX
XX PD 15-DEC-1992.
XX XX
XX XX 06-NOV-1991; 91US-00788262.
XX PF 30-SEP-1983; 88US-00536911.
XX PR 27-SEP-1984; 84US-00654942.
XX PR 06-OCT-1988; 88US-00256135.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
XX PI Racanelliello VR, Baroudy BM, Emerson SU;
XX PI WPI; 1993-067429/08.
XX DR N-PSDB; AAQ36934.
XX CC Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of
XX FT antigen and antibodies.
XX PS Disclosure; Fig 7; 65pp; English.
XX CC HAV virion RNA was extracted from the livers of marmosets which had been
XX CC inoculated with HAV (the HAV had previously been passaged twice in
XX CC marmosets). The RNA was used to prepare ds cDNA clones by standard
XX CC methods. Clones contg. inserts which hybridised to RNA from HAV-infected
XX CC African Green Monkey kidney cells were selected for further analysis. A
XX CC 7.4kb restriction map (about 99% of the HAV genome) was constructed from
XX CC 5 overlapping inserts. The sequence of the first 3.3kb (approx.) from the
XX CC 5'-terminus was determined. An amino acid sequence was deduced from the
XX CC entire clone and an open reading frame was identified starting at
XX CC position 238. A comparison of the predicted HAV amino acid sequences with
XX CC the known capsid protein sequences of other picornaviruses (poliovirus,
XX CC foot and mouth disease virus and encephalomyelitis virus) revealed areas
XX CC of local homology. (Note: Revised entry submitted to correct the patent
XX CC number format of US Government-owned NTIS applications to prevent clashes
XX CC with ongoing US granted patent numbers. For further information please
XX CC visit the Derwent web site at www.derwent.com/dwpi/updates/ntis us.html.)
XX CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to
XX CC correct OS field.)
XX CC
XX CC Sequence 1091 AA;
XX CC
Query Match 100.0%; Score 101; DB 2; Length 1091;
Best Local Similarity 100.0%; Pred. No. 9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSOAKISLFYTEHEIMKFS 20
Db 1071 FSOAKISLFYTEHEIMKFS 1090
RESULT 6
AAR05697
ID AAR05697 standard; protein; 2227 AA.
XX AC AAR05697;
XX DT 24-OCT-2003 (revised)
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 15-AUG-1990 (first entry)
XX CC Attenuated hepatitis A virus.
XX CC
XX KW Hepatitis A virus; vaccine; attenuated.
XX OS Hepatitis A virus; strain HM-175.
XX FH Key Location/Qualifiers
XX FT Region 1. .23
XX FT /label= VP4 = 1A
XX FT 24. .245
XX FT /label= VP2 = 1B
XX FT 246. .491
XX FT /label= VP3 = 1C
XX FT 492. .791
XX FT /label= VP1 = 1D
XX FT 792. .980
XX FT /label= 2A
XX FT 981. .1087

FT Region /label= 2B
FT 1088..1422
FT /label= 2C
FT 1423..1496
FT /label= 3A
FT 1497..1519
FT /label= 3B = VPg
FT 1520..1738
FT /label= 3C
FT 1739..2227
FT /label= 3D
XX
XX US4894228-A.
XX
XX 16-JAN-1990.
XX
XX 12-JUL-1988; 88US-00217824.
XX
XX 19-SEP-1984; 84US-00652067.
XX 09-SEP-1986; 86US-00905146.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
XX Daemer RJ, Gust ID;
XX
XX WPI; 1990-075557/10.
XX N-PSDB; AAQ03512.
XX
XX Vaccine against hepatitis A virus infection - comprises novel attenuated
XX hepatitis A virus strain.
XX
XX Claim 1; Fig 1; 18pp; English.
XX
XX The attenuated HAV is useful for inducing protective immunity against
XX HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
XX several nucleotide changes distributed throughout the genome, is
XX attenuated for chimpanzees, elicits serum neutralising antibodies, and is
XX suitable for use as an HAV vaccine. It is noted that not all the changes
XX are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
XX 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
XX standardise OS field)
XX Sequence 2227 AA;
XX
XX Query Match 100.0%; Score 101; DB 2; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 2e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FSQAKISLFYTEHEIMKFS 20
XX |||||||||||||||||||||
XX 834 FSQAKISLFYTEHEIMKFS 853
XX
XX RESULT 7
XX AAW34074
XX ID AAW34074 standard; protein; 2227 AA.
XX AC AAW34074;
XX
XX 17-OCT-2003 (revised)
XX 27-APR-1998 (first entry)
XX
XX Hepatitis A virus HM-175 protein sequence.
XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
XX vaccine.
XX
XX Hepatitis A virus; HM-175.
XX OS
XX FH Key Location/Qualifiers
XX Protein 1..23

FT Protein /label= VP4
FT 24..245
FT /label= VP2
FT 246..491
FT /label= VP3
FT 492..791
FT /label= VP1
FT 792..980
FT /label= 2A
FT 981..1087
FT /label= 2B
FT 1088..1422
FT /label= 2C
FT 1423..1496
FT /label= 3A
FT 1497..1519
FT /label= 3B
FT 1520..1738
FT /label= 3C
FT 1739..2227
FT /label= 3D
XX
XX WO9740166-A2.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US006506.
XX
XX 19-APR-1996; 96US-0015642P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Raychaudhuri G, Emerson SU, Purcell RH;
XX
XX WPI; 1997-535850/49.
XX N-PSDB; AAT93023.
XX
XX Human attenuated HAV genome containing simian HAV 2C gene - useful as
XX vaccines against HAV infection.
XX
XX Disclosure; Fig 13A-D; 66pp; English.
XX
XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
XX 175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
XX by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
XX construct (I) comprises a genome of HAV, where the genome is a human
XX attenuated HAV genome in which a region of the 2C gene has been replaced
XX by a corresponding region from a 2C gene of a simian AGM-27 HAV genome
XX (see AAT93024). The region of the 2C gene from AGM-27 contained in the
XX construct preferably encodes amino acids 120-328 of the 2C protein, amino
XX acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript
XX of (I); (2) a cell transfected with (I) or the RNA transcript of (1); (3)
XX a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host
XX cell containing the HAV of (3). (I) or its RNA transcript, can be used as
XX a vaccine for preventing HAV in a mammal. (I) or the RNA transcript can
XX also be used to stimulate the production of protective antibodies in the
XX mammal. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 2227 AA;
XX
XX Query Match 100.0%; Score 101; DB 2; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 2e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FSQAKISLFYTEHEIMKFS 20
XX |||||||||||||||||||||
XX 834 FSQAKISLFYTEHEIMKFS 853
XX
XX RESULT 8
XX AAB18609
XX ID AAB18609 standard; protein; 2227 AA.
XX
XX

AC AAB18609;
 XX DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW HAV 4380.
 XX
 OS Hepatitis A virus.
 XX
 PN US6113912-A.
 XX PD 05-SEP-2000.
 XX PF 07-JUN-1995; 95US-00475886.
 XX PR 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX WPI; 2000-586464/55.
 DR N-PSDB; AAA75478.
 XX
 PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
 line useful as vaccine for protecting humans against hepatitis A virus
 infection, has modified genome compared to wild type.
 XX
 PS Disclosure; Fig 6A-K; 72pp; English.
 XX
 CC The present sequence is derived from a wild type hepatitis A virus (HAV)
 strain HM-174. The sequence is modified to produce HAV which are adapted
 CC to growth in the human fibroblast-like cell line MRC-5. The HAV is able
 CC to propagate in MRC-5 cells and retain appropriate attenuation. It is
 CC useful as a live vaccine for prophylaxis of hepatitis A in humans and
 CC other primates
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 3; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2e-07; 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0;
 QY 1 FSOAKISLFYTEHEIMKFS 20
 |||||
 DB 834 FSOAKISLFYTEHEIMKFS 853
 RESULT 10
 AAB18608
 ID AAB18608 standard; protein; 2227 AA.
 XX
 AC AAB18608;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW P-35 virus.
 XX
 OS Hepatitis A virus.
 XX
 PN US6113912-A.
 XX PD 05-SEP-2000.
 XX PF 07-JUN-1995; 95US-00475886.
 XX PR 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX WPI; 2000-586464/55.
 DR N-PSDB; AAA75477.
 XX
 PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
 line useful as vaccine for protecting humans against hepatitis A virus
 infection, has modified genome compared to wild type.
 XX
 PS Disclosure; Col 67-78; 72pp; English.
 XX

AC AAB18609;
 XX DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW HAV 4380.
 XX
 OS Hepatitis A virus.
 XX
 PN US6113912-A.
 XX PD 05-SEP-2000.
 XX PF 07-JUN-1995; 95US-00475886.
 XX PR 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX WPI; 2000-586464/55.
 DR N-PSDB; AAA75478.
 XX
 PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
 line useful as vaccine for protecting humans against hepatitis A virus
 infection, has modified genome compared to wild type.
 XX
 PS Disclosure; Col 93-104; 72pp; English.
 XX
 CC The present sequence is derived from a live attenuated hepatitis A virus
 (HAV) of the invention, designated HAV 4380. The sequence is produced by
 CC modifying wild type HAV strain HM-174. The HAV of the invention are
 CC adapted to growth in the human fibroblast-like cell line MRC-5. The HAV
 CC is able to propagate in MRC-5 cells and retain appropriate attenuation.
 CC It is useful as a live vaccine for prophylaxis of hepatitis A in humans
 CC and other primates
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 3; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSOAKISLFYTEHEIMKFS 20
 |||||
 DB 834 FSOAKISLFYTEHEIMKFS 853
 RESULT 9
 AAB18607
 ID AAB18607 standard; protein; 2227 AA.
 XX
 AC AAB18607;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
 XX
 OS Hepatitis A virus.
 XX
 PN US6113912-A.
 XX PD 05-SEP-2000.
 XX PF 07-JUN-1995; 95US-00475886.
 XX

CC The present sequence is derived from passage 35 of a wild type hepatitis
CC A virus (HAV) strain HM-174. The resulting virus is designated P-35
CC virus. The sequence is modified to produce HAV which are adapted to
CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
CC as a live vaccine for prophylaxis of hepatitis A in humans and other
CC primates
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
DB 834 FSOAKISLFYTEEHEIMKFS 853

RESULT 11
AAE19899
ID AAE19899 standard; protein; 2227 AA.

XX AC AAE19899;
XX DT 18-JUN-2002 (first entry)
XX DE Hepatitis A virus (HAV) protein.
XX KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX KW cytosstatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX OS Hepatitis A virus.
XX PN W0200213855-A2.
XX PD 21-FEB-2002.

XX PF 15-AUG-2001; 2001WO-1B001808.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.

XX PA (TRIP-) TRIPEP AB.

XX PI Sallberg M, Hultgren C;

XX DR WPI; 2002-241837/29.

XX DR N-PSDB; AAD31766.

XX PT Vaccine compositions for treating and preventing disease, preferably
XX PT hepatitis C virus infection, comprises ribavirin and antigen that has
XX PT epitope present in hepatitis C virus.

XX PS Claim 11; Page 82-87; 120pp; English.

XX CC The invention relates to a composition comprising ribavirin and an
XX CC antigen preferably non structural 3 protein (NS3)/4A fragment of
XX CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX CC sequence. The composition is useful for enhancing an immune response to a
XX CC hepatitis C antigen in humans, domestic, sport or pet species and as
XX CC vaccines for treating and preventing HCV infections. The composition is
XX CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX CC The present sequence is hepatitis A virus (HAV) protein

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20

DB 834 FSOAKISLFYTEEHEIMKFS 853
RESULT 12
ABG31729
ID ABG31729 standard; protein; 2227 AA.
XX AC ABG31729;
XX DT 29-AUG-2003 (revised)
XX DT 29-NOV-2002 (first entry)
XX DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
XX KW HAV 4380.
XX OS Hepatitis A virus; strain HM-175.
XX PN US6423318-B1.
XX PD 23-JUL-2002.
XX PF 31-AUG-2000; 2000US-00653499.
XX PR 17-SEP-1993; 93WO-US008610.
XX PR 17-APR-1995; 95US-00397232.
XX PR 07-JUN-1995; 95US-00475886.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX DR WPI; 2002-680946/73.
XX DR N-PSDB; ABS52789.

XX PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX PS Disclosure; Col 93-104; 71pp; English.
XX CC The invention relates to a polynucleotide which encodes a hepatitis A
XX CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX CC line). The polynucleotide is useful for preparing a vaccine against
XX CC hepatitis A virus infection. This sequence represents an attenuated
XX CC hepatitis A virus 4380 polypeptide. (Updated on 29-AUG-2003 to
XX CC standardise OS field)
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
DB 834 FSOAKISLFYTEEHEIMKFS 853

RESULT 13
ABG31727

ID ABG31727 standard; protein; 2227 AA.

XX AC ABG31727;

XX DT 29-AUG-2003 (revised)

XX DT 29-NOV-2002 (first entry)

XX DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.

XX KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.

```

XX Hepatitis A virus; strain HM-175.
OS US6423318-B1.
XX
PN 23-JUL-2002.
XX
PD 31-AUG-2000; 2000US-00653499.
XX
PF 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2002-680946/73.
DR N-PSDB; ABS52787.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX
XX Disclosure; Fig 6; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
CC field)
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 101; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853
RESULT 14
ABG31728
ID ABG31728 standard; protein; 2227 AA.
XX
XX ABG31728;
AC 29-NOV-2002 (first entry)
XX
XX Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.
DE
XX Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
KW virucide; mutant; pHAV/7; mutain.
KW
XX Hepatitis A virus; strain HM-175.
OS Synthetic.
OS
XX Key Location/Qualifiers
FH Misc-difference 764
FT Misc-difference 821 /note= "Wild-type Glu substituted by Val"
FT Misc-difference 821 /note= "Wild-type Asn substituted by Ser"
FT Misc-difference 963 /label= Wild-type Lys substituted by Arg
FT Misc-difference 1052 /note= "Wild-type Ala substituted by Val"
FT Misc-difference 1062 /note= "Wild-type Gly substituted by Ala"
FT Misc-difference 1118 /note= "Wild-type Lys substituted by Met"
FT

```

```

FT Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"
FT Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"
FT Misc-difference 1277 /note= "Wild-type Val substituted by Ile"
FT Misc-difference 1500 /note= "Wild-type His substituted by Tyr"
FT Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"
XX
XX US6423318-B1.
XX
XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-00653499.
XX
XX 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2002-680946/73.
DR N-PSDB; ABS52788.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX
XX Example 3; Col 67-78; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC mutant strain HM-175/7 (pHAV/7) polypeptide
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 101; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853
RESULT 15
ABU08640
ID ABU08640 standard; protein; 2227 AA.
XX
XX ABU08640;
AC 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
XX Attenuated (pass35) hepatitis A virus strain HM-175.
XX
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
XX Hepatitis A virus; strain HM-175.
OS
XX US2002176869-A1.
PN 28-NOV-2002.
XX

```

PF 29-APR-2002; 2002US-00135988.
XX
PR 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
PA (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX
DR WPI; 2003-352605/02.
DR N-PSDB; ABX93474.
XX
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
PS Example 3; Fig 6; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated (pass 35)
CC human hepatitis A virus strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853
Search completed: March 15, 2004, 13:59:57
Job time : 45.2353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:15:47 ; Search time 11.2941 Seconds
(without alignments)
91.421 Million cell updates/sec

Title: US-09-171-432A-43
Perfect score: 101
Sequence: 1 FSOAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pdp:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pdp:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pdp:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pdp:*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pdp:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1091	6	Patent No. 5516630
2	101	100.0	2227	3	US-08-475-886-2
3	101	100.0	2227	3	US-08-475-886-4
4	101	100.0	2227	3	US-08-475-886-6
5	101	100.0	2227	3	US-08-397-232-2
6	101	100.0	2227	3	US-08-397-232-4
7	101	100.0	2227	3	US-09-171-387-2
8	101	100.0	2227	4	US-09-653-499-2
9	101	100.0	2227	4	US-09-653-499-4
10	101	100.0	2227	4	US-09-653-499-6
11	101	100.0	2227	4	US-10-104-966-12
12	101	100.0	2227	4	US-10-135-988-2
13	101	100.0	2227	4	US-10-135-988-4
14	101	100.0	2227	4	US-10-135-988-6
15	46	45.5	741	4	US-09-489-039A-8111
16	44	43.6	347	4	US-09-543-681A-6173
17	43	42.6	171	4	US-09-134-001C-5205
18	41	40.6	165	4	US-09-107-532A-7030
19	41	40.6	173	4	US-09-585-228-4
20	41	40.6	2431	1	US-07-920-281C-2
21	41	40.6	2431	3	US-08-466-277-2
22	40	39.6	173	4	US-09-585-228-2
23	40	39.6	240	4	US-09-328-352-7537
24	40	39.6	243	4	US-09-252-991A-17307
25	39	38.6	148	4	US-09-198-452A-491
26	39	38.6	170	4	US-09-198-119C-89
27	39	38.6	289	4	US-09-305-856B-8

```

28 39 38.6 289 5 PCT-US92-00282-11 Sequence 11, Appl
29 39 38.6 340 4 US-09-107-532A-3720 Sequence 3720, Ap
30 39 38.6 455 4 US-09-198-452A-372 Sequence 372, App
31 39 38.6 530 4 US-09-252-991A-23666 Sequence 23666, A
32 39 38.6 668 4 US-09-543-681A-7603 Sequence 21541, A
33 39 38.6 715 4 US-09-543-681A-7603 Sequence 7603, Ap
34 39 38.6 860 1 US-08-092-817-4 Sequence 4, Appli
35 39 38.6 860 4 US-08-485-128-4 Sequence 4, Appli
36 39 38.6 860 4 US-09-804-778A-8 Sequence 8, Appli
37 39 38.6 1410 2 US-08-470-058-4 Sequence 4, Appli
38 39 38.6 1410 3 US-09-037-188-4 Sequence 4, Appli
39 39 38.6 1410 3 US-09-285-310-4 Sequence 4, Appli
40 38 37.6 265 1 US-07-958-551-2 Sequence 2, Appli
41 38 37.6 265 1 US-08-129-610-7 Sequence 7, Appli
42 38 37.6 265 1 US-08-129-609A-7 Sequence 7, Appli
43 38 37.6 265 1 US-08-455-313-7 Sequence 7, Appli
44 38 37.6 265 1 US-08-475-924-2 Sequence 2, Appli
45 38 37.6 265 2 US-08-657-579A-2 Sequence 2, Appli

```

ALIGNMENTS

```

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.;
; BAROUY, BAIGIE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2
; LENGTH: 1091
5516630-2

```

```

Query Match 100.0%; Score 101; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FSOAKISLFYTEHEIMKFS 20
Db 1071 FSOAKISLFYTEHEIMKFS 1090

```

```

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1

```

```
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEEHEIMKFS 20
Db 834 FSQAKISLFYTEEHEIMKFS 853

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEEHEIMKFS 20
Db 834 FSQAKISLFYTEEHEIMKFS 853

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEEHEIMKFS 20
Db 834 FSQAKISLFYTEEHEIMKFS 853

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEEHEIMKFS 20
Db 834 FSQAKISLFYTEEHEIMKFS 853

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 7
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.;
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 8
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 9
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 10
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 7
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.;
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 8
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 9
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 10
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 7
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.;
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 8
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 9
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853

RESULT 10
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31

```

; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
; US-09-653-499-6

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSQAKISLFYTEHEIMKFS 20
DB      834 FSQAKISLFYTEHEIMKFS 853

RESULT 11
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Saliberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSQAKISLFYTEHEIMKFS 20
DB      834 FSQAKISLFYTEHEIMKFS 853

RESULT 12
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; PRIOR FILING DATE: 1992-09-18

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSQAKISLFYTEHEIMKFS 20
DB      834 FSQAKISLFYTEHEIMKFS 853

RESULT 13
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSQAKISLFYTEHEIMKFS 20
DB      834 FSQAKISLFYTEHEIMKFS 853

RESULT 14
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:24:17 ; Search time 55.2941 Seconds
(without alignments)
127.748 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRLGLSGVGEIKEQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	25	2	Aaw42969 Immunogen
2	129	100.0	25	4	Aab69448 Synthetic
3	129	100.0	2227	1	Aap60066 Sequence
4	129	100.0	2227	2	Aar05697 Attenuate
5	129	100.0	2227	2	Aaw34074 Hepatitis
6	129	100.0	2227	3	Aab18609 Amino aci
7	129	100.0	2227	3	Aab18607 Amino aci
8	129	100.0	2227	3	Aab18608 Amino aci
9	129	100.0	2227	5	Aae19899 Hepatitis
10	129	100.0	2227	5	Abg31729 Attenuate
11	129	100.0	2227	5	Abg31727 Wild-type
12	129	100.0	2227	5	Abg31728 Hepatitis
13	129	100.0	2227	6	Abu08640 Attenuat
14	129	100.0	2227	6	Abu08641 Attenuat
15	129	100.0	2227	6	Abu08639 Wild type
16	129	100.0	2227	7	Abw00350 Hepatitis
17	96	74.4	20	2	Aaw42929 Immunogen
18	96	74.4	20	4	Aab59446 Synthetic
19	49	38.0	229	2	Aaw32621 Bacillus
20	49	38.0	474	2	AAR0507 S. livida
21	48.5	37.6	1987	7	Abu63358 Human Sin
22	48.5	37.6	2013	7	Abu63356 Human Sin
23	48.5	37.6	2014	4	Ag67395 Amino aci
24	48.5	37.6	2014	5	Aae24137 Human kin
25	48.5	37.6	2014	6	Abp60434 Human ste

26	48.5	37.6	2014	7	ABU63359	Human Sin
27	48.5	37.6	2040	7	ABU63357	Human Sin
28	48.5	37.6	2041	6	ABP60435	Human ste
29	48	37.2	213	4	ABB59304	Drosophil
30	48	37.2	430	7	ADB70157	C. neofor
31	48	37.2	841	6	ABU23488	Protein e
32	47.5	36.8	603	6	ABU20315	Protein e
33	46.5	36.0	146	6	ABU19172	Protein e
34	46.5	36.0	993	4	ABB66196	Drosophil
35	46	35.7	342	6	ABU26593	Protein e
36	46	35.7	877	3	AAV96168	Saccharom
37	45.5	35.3	1066	6	ABU19620	Protein e
38	45	34.9	140	4	ABG27711	Novel hum
39	45	34.9	140	7	ADC33299	Human nov
40	45	34.9	177	4	AAAB36623	Human FLE
41	45	34.9	177	7	ADC31745	Human nov
42	45	34.9	308	4	AAAB8771	Human T2R
43	45	34.9	317	7	ADD01168	Human nuc
44	45	34.9	325	4	AAU16582	Human nov
45	45	34.9	325	6	ABU55651	Human nov

ALIGNMENTS

RESULT 1
AAW42969
ID AAW42969 standard; peptide; 25 AA.
XX
AC AAW42969;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1757.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC The present immunogenic peptide corresponds to an immunogenic epitope of
CC the Hepatitis A virus (HAV). The peptide is substantially similar to a
CC portion of the amino acid sequence of the P2A protein of HAV
CC corresponding to amino acids 792-980. Compositions containing the peptide
CC can be used to induce an immune response to HAV in a mammal. The peptide
CC can also be used to detect the presence of antibodies against HAV in
CC mammalian serum. The peptide can also be used to make an antibody against
CC HAV by administering the peptide to a mammal
XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 129; DB 2; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
|||||
Db 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25

RESULT 2

AAB69448

ID AAB69448 standard; peptide; 25 AA.

XX AC AAB69448;

XX DT 20-APR-2001 (first entry)

XX XX Synthetic HAV P2A peptide, SEQ ID NO: 48.

XX DE Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
XX KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX OS Hepatitis A virus.
XX OS Synthetic.
XX PN WO200105824-A2.
XX PD 25-JAN-2001.
XX PF 14-JUL-2000; 2000WO-US019267.
XX PR 15-JUL-1999; 99US-0144412P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;
XX XX WPI; 2001-112691/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays
XX PT detecting anti-hepatitis A virus and as vaccines.
XX PS Claim 13; Page 99; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity

XX Sequence 25 AA;

Query Match 100.0%; Score 129; DB 4; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.9e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25

Db 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25

RESULT 3

AAP60066

ID AAP60066 standard; protein; 2227 AA.

XX AAP60066;

XX AC 25-MAR-2003 (revised)

XX DT 26-JUN-1991 (first entry)

XX DE Sequence of viral L434 polypeptide encoded by the complete nucleotide
XX DE sequence of the HAV genome.

XX KW Diagnosis; vaccine; passive immunotherapy.

XX OS Hepatitis A virus.

XX FH Key Location/Qualifiers

XX FT Region 1. .245

XX FT Region /label= P1.1A

XX FT Region 246. .491

XX FT Region /label= 1B

XX FT Region 492. .836

XX FT Region /label= 1C

XX FT Region 837. .980

XX FT Region /label= P2.2A

XX FT Region 981. .1076

XX FT Region /label= 2B

XX FT Region 1077. .1422

XX FT Region /label= 2C

XX FT Region 1423. .1484

XX FT Region /label= P3.3A

XX FT Region 1485. .1507

XX FT Region /label= 3B

XX FT Region 1508. .1678

XX FT Region /label= 3C

XX FT Region 1679. .2227

XX FT Region /label= 3D

XX PN EPI99480-A.

XX PD 29-OCT-1986.

XX PF 03-APR-1986; 86EP-00302465.

XX PR 03-APR-1985; 85US-00719329.

XX PA (CHIR) CHIRON CORP.

XX PI Dina D, Potter SJ, Vannest GA, Caput D;

XX DR WPI; 1986-286213/44.

XX DR N-PSDB; AAN60080.

XX PT Hepatitis A virus nucleotide sequence and polypeptide - and use in prodn.

XX PS of vaccines and diagnostic probes.

XX PS Claim 5; Fig 1; 18pp; English.

XX CC AAN60080 and oligonucleotide fragments are useful in detection of

CC hepatitis A virus; transformed hosts may be used for expression of

CC polypeptides and fragments useful in vaccines without risk of infection

CC by the virus or in prodn. of particles which are capable of inducing

CC immunocompetent B cells for passive immunotherapy. Pref. epitope is

CC derived from AAs 445-657 or 792-848 of the HAV polypeptide sequence

CC (AAP60066). (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.1e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25

Db 956 WLNPKKINLADRLMLGSLGVQEIKEQ 980

CC standardise OS field)

XX Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 2; Length 2227;

Best Local Similarity 100.0%; Pred. NO. 4,1e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKINLADRMGLSGVQEIKEQ 25

Db 956 WLNPKINLADRMGLSGVQEIKEQ 980

RESULT 5

AAW34074

ID AAW34074 standard; protein; 2227 AA.

XX AC AAW34074;

XX DT 17-OCT-2003 (revised)

XX DT 27-APR-1998 (first entry)

XX DE Hepatitis A virus HM-175 protein sequence.

XX KW HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;

XX KW vaccine.

XX OS Hepatitis A virus; HM-175.

XX FH Key Location/Qualifiers

FT Protein 1..23

FT Protein /label= VP4

FT Protein 24..245

FT Protein /label= VP2

FT Protein 246..491

FT Protein /label= VP3

FT Protein 492..791

FT Protein /label= VP1

FT Protein 792..980

FT Protein /label= 2A

FT Protein 981..1087

FT Protein /label= 2B

FT Protein 1088..1422

FT Protein /label= 2C

FT Protein 1423..1496

FT Protein /label= 3A

FT Protein 1497..1519

FT Protein /label= 3B

FT Protein 1520..1738

FT Protein /label= 3C

FT Protein 1739..2227

FT Protein /label= 3D

XX PN WO9740166-A2.

XX PD 30-OCT-1997.

XX PF 18-APR-1997; 97WO-US006506.

XX PR 19-APR-1996; 96US-0015642P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Raychaudhuri G, Emerson SU, Purcell RH;

XX DR WPI; 1997-535850/49.

XX DR N-PSDB; AAT93023.

XX PT Human attenuated HAV genome containing simian HAV 2C gene - useful as

XX PT vaccines against HAV infection.

XX PS Disclosure; Fig 13A-D; 66pp; English.

RESULT 4

AAW05697

ID AAR05697 standard; protein; 2227 AA.

XX AC AAR05697;

XX DT 24-OCT-2003 (revised)

XX DT 27-AUG-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 15-AUG-1990 (first entry)

XX DE Attenuated hepatitis A virus.

XX KW Hepatitis A virus; vaccine; attenuated.

XX OS Hepatitis A virus; strain HM-175.

XX FH Key Location/Qualifiers

FT Region 1..23

FT Region /label= VP4 = 1A

FT Region 24..245

FT Region /label= VP2 = 1B

FT Region 246..491

FT Region /label= VP3 = 1C

FT Region 492..791

FT Region /label= VP1 = 1D

FT Region 792..980

FT Region /label= 2A

FT Region 981..1087

FT Region /label= 2B

FT Region 1088..1422

FT Region /label= 2C

FT Region 1423..1496

FT Region /label= 3A

FT Region 1497..1519

FT Region /label= 3B = VPg

FT Region 1520..1738

FT Region /label= 3C

FT Region 1739..2227

FT Region /label= 3D

XX PN US4894228-A.

XX PD 16-JAN-1990.

XX PF 12-JUL-1988; 88US-00217824.

XX PR 19-SEP-1984; 84US-00652067.

XX PR 09-SEP-1986; 86US-00905146.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

XX PI Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstone SM;

XX PI Daemer RJ, Gust ID;

XX WPI; 1990-075557/10.

XX DR N-PSDB; AAQ03512.

XX PT Vaccine against hepatitis A virus infection - comprises novel attenuated

XX PT hepatitis A virus strain.

XX PS Claim 1; Fig 1; 18pp; English.

XX CC The attenuated HAV is useful for inducing protective immunity against

XX CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by

XX CC several nucleotide changes distributed throughout the genome, is

XX CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is

XX CC suitable for use as an HAV vaccine. It is noted that not all the changes

XX CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-

XX CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

XX CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to

CC This protein sequence is encoded by the human hepatitis A virus (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA construct (1) comprises a genome of HAV, where the genome is a human attenuated HAV genome in which a region of the 2C gene has been replaced by a corresponding region from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The region of the 2C gene from AGM-27 contained in the construct preferably encodes amino acids 120-328 of the 2C protein, amino acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (1) or its RNA transcript, can be used as a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can also be used to stimulate the production of protective antibodies in the mammal. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 6
AAB18609
ID AAB18609 standard; protein; 2227 AA.
XX AAB18609;
XX
XX 15-JAN-2001 (first entry)
XX
XX Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW HAV 4380.
XX Hepatitis A virus.
XX
XX US6113912-A.

XX 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-00475886.
XX
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI: 2000-586464/55.
XX N-PSDB; AAA75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.
XX Disclosure; Col 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A virus (HAV) of the invention, designated HAV 4380. The sequence is produced by modifying wild type HAV strain HM-174. The HAV of the invention are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates

SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 7
AAB18607
ID AAB18607 standard; protein; 2227 AA.
XX
XX AAB18607;
XX
XX 15-JAN-2001 (first entry)

XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX Hepatitis A virus.
XX US6113912-A.

XX 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-00475886.
XX
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI: 2000-586464/55.
XX N-PSDB; AAA75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.
XX Disclosure; Fig 6A-K; 72pp; English.

XX The present sequence is derived from a wild type hepatitis A virus (HAV) strain HM-174. The sequence is modified to produce HAV which are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates

SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 8
AAB18608
ID AAB18608 standard; protein; 2227 AA.
XX AAB18608;
XX
XX 15-JAN-2001 (first entry)

XX DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW P-35 virus.
XX Hepatitis A virus.
XX US6113912-A.
XX PD 05-SEP-2000.
XX PF 07-JUN-1995; 95US-00475886.
XX PR 18-SEP-1992; 92US-00947338.
XX PR 17-SEP-1993; 93WO-US008610.
XX PR 17-APR-1995; 95US-00397232.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.
XX N-PSDB; AAA75477.
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type.
XX Disclosure; Col 67-78; 72pp; English.
XX The present sequence is derived from passage 35 of a wild type hepatitis
CC A virus (HAV) strain HM-174. The resulting virus is designated P-35
CC virus. The sequence is modified to produce HAV which are adapted to
CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
CC as a live vaccine for prophylaxis of hepatitis A in humans and other
CC primates
XX
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKINLADRLMLGLSGVQEIKEQ 25
DB 956 WLNPKINLADRLMLGLSGVQEIKEQ 980
RESULT 9
AAE19899
ID AAE19899 standard; protein; 2227 AA.
AC AAE19899;
XX 18-JUN-2002 (first entry)
XX Hepatitis A virus (HAV) protein.
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX Hepatitis A virus.
OS WO200213855-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-18001808.
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
PR

PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
DR N-PSDB; AAD31766.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Claim 11; Page 82-87; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is hepatitis A virus (HAV) protein
XX
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKINLADRLMLGLSGVQEIKEQ 25
DB 956 WLNPKINLADRLMLGLSGVQEIKEQ 980
RESULT 10
ABG31729
ID ABG31729 standard; protein; 2227 AA.
XX AC ABG31729;
XX 29-AUG-2003 (revised)
DT 29-NOV-2002 (first entry)
XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
DE Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
KW HAV 4380.
XX Hepatitis A virus; strain HM-175.
OS US6423318-B1.
XX 23-JUL-2002.
PD 31-AUG-2000; 2000US-00653499.
XX 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2002-680946/73.
DR N-PSDB; ABS52789.
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX Disclosure; Col 93-104; 71pp; English.
PS

XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents an attenuated
CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980
|||||

RESULT 11
ABG31727
ID ABG31727 standard; protein; 2227 AA.
XX
AC ABG31727;
XX
DT 29-AUG-2003 (revised)
DT 29-NOV-2002 (first entry)
XX
DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
XX
KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
XX
OS Hepatitis A virus; strain HM-175.
XX
PN US6423318-B1.
XX
PD 23-JUL-2002.
XX
PF 31-AUG-2000; 2000US-00653499.
XX
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX WPI; 2002-680946/73.
XX N-PSDB; ABS52787.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX
XX Disclosure; Fig 6; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
CC field)
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980
|||||

RESULT 12
ABG31728
ID ABG31728 standard; protein; 2227 AA.
XX
AC ABG31728;
XX
DT 29-NOV-2002 (first entry)
XX
DE Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
KW virucide; mutant; pHAV/7; mutcin.
XX
OS Hepatitis A virus; strain HM-175.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 764 /note= "Wild-type Glu substituted by Val"
FT Misc-difference 821 /note= "Wild-type Asn substituted by Ser"
FT Misc-difference 963 /label= Wild-type Lys substituted by Arg
FT Misc-difference 1052 /note= "Wild-type Ala substituted by Val"
FT Misc-difference 1062 /note= "Wild-type Gly substituted by Ala"
FT Misc-difference 1118 /note= "Wild-type Lys substituted by Met"
FT Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"
FT Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"
FT Misc-difference 1277 /note= "Wild-type Val substituted by Ile"
FT Misc-difference 1500 /note= "Wild-type His substituted by Tyr"
FT Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"
XX
XX US6423318-B1.
XX
XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-00653499.
XX
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX 07-JUN-1995; 95US-00475886.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX WPI; 2002-680946/73.
XX N-PSDB; ABS52788.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX
XX Example 3; Col 67-78; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC mutant strain HM-175/7 (pHAV/7) polypeptide
XX


```
SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
   |||||
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 13
ABU08640
ID ABU08640 standard; protein; 2227 AA.
XX AC ABU08640;
XX DT 23-OCT-2003 (revised)
XX DT 03-JUN-2003 (first entry)
XX DE Attenuated (pass35) hepatitis A virus strain HM-175.
XX KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
XX KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX OS Hepatitis A virus; strain HM-175.
XX PN US2002176869-A1.
XX PD 28-NOV-2002.
XX PF 29-APR-2002; 2002US-00135988.
XX PR 18-SEP-1992; 92US-00947338.
XX PR 17-SEP-1993; 93WO-US008610.
XX PR 17-APR-1995; 95US-00397232.
XX PR 07-JUN-1995; 95US-00475886.
XX PR 31-AUG-2000; 2000US-00653499.
XX PA (FUNK/) FUNKHOUSER A W.
XX PA (EMER/) EMERSON S U.
XX PA (FURC/) PURCELL R H.
XX PA (DHON/) D'HONDT E.
XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2003-352605/02.
XX DR N-PSDB; ABX93474.
XX PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
XX PT useful in vaccines for protecting primates against hepatitis infection
XX PT and disease.
XX PS Disclosure; Page 45-51; 70pp; English.
XX CC The invention describes a live hepatitis A virus (HAV) adapted to growth
XX CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
XX CC a vaccine for protecting primates against hepatitis infection and
XX CC disease. This is the amino acid sequence of an attenuated human
XX CC hepatitis A virus (4380) strain HM-175. (Updated on 23-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
   |||||
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 14
ABU08641
ID ABU08641 standard; protein; 2227 AA.
XX AC ABU08641;
XX DT 23-OCT-2003 (revised)
XX DT 03-JUN-2003 (first entry)
XX DE Attenuated hepatitis A virus (4380) strain HM-175.
XX KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
XX KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX OS Hepatitis A virus; strain HM-175.
XX PN US2002176869-A1.
XX PD 28-NOV-2002.
XX PF 29-APR-2002; 2002US-00135988.
XX PR 18-SEP-1992; 92US-00947338.
XX PR 17-SEP-1993; 93WO-US008610.
XX PR 17-APR-1995; 95US-00397232.
XX PR 07-JUN-1995; 95US-00475886.
XX PR 31-AUG-2000; 2000US-00653499.
XX PA (FUNK/) FUNKHOUSER A W.
XX PA (EMER/) EMERSON S U.
XX PA (FURC/) PURCELL R H.
XX PA (DHON/) D'HONDT E.
XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2003-352605/02.
XX DR N-PSDB; ABX93475.
XX PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
XX PT useful in vaccines for protecting primates against hepatitis infection
XX PT and disease.
XX PS Disclosure; Page 45-51; 70pp; English.
XX CC The invention describes a live hepatitis A virus (HAV) adapted to growth
XX CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
XX CC a vaccine for protecting primates against hepatitis infection and
XX CC disease. This is the amino acid sequence of an attenuated human
XX CC hepatitis A virus (4380) strain HM-175. (Updated on 23-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
   |||||
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 15
ABU08639
ID ABU08639 standard; protein; 2227 AA.
XX AC ABU08639;
XX DT 23-OCT-2003 (revised)
XX DT 03-JUN-2003 (first entry)
XX DE Wild type human hepatitis A virus strain HM-175.
XX
```

KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection.
XX
OS Hepatitis A virus; strain HM-175.
XX

XX US2002176869-A1.

XX PD 28-NOV-2002.

XX PF 29-APR-2002; 2002US-00135988.

XX PR 18-SEP-1992; 92US-00947338.

XX PR 17-SEP-1993; 93WO-US008610.

XX PR 17-APR-1995; 95US-00397232.

XX PR 07-JUN-1995; 95US-00475886.

XX PR 31-AUG-2000; 2000US-00653499.

XX PA (FUNK/) FUNKHOUSER A W.

XX PA (EMER/) EMERSON S U.

XX PA (PURC/) PURCELL R H.

XX PA (DHON/) D'HONDT E.

XX FI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX DR WPI; 2003-352605/02.

XX DR N-PSDB; ABX93473.

XX PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
XX useful in vaccines for protecting primates against hepatitis infection
XX and disease.

XX FS Disclosure; Fig 6; 70pp; English.

XX CC The invention describes a live hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
XX a vaccine for protecting primates against hepatitis infection and
XX disease. This is the amino acid sequence of wild type human hepatitis A
XX virus strain HM-175. (Updated on 23-OCT-2003 to standardise OS field)

SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 6; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.1e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVQEIKEQ 25

DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

Search completed: March 15, 2004, 14:00:01

Job time : 56.2941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:15:47 ; Search time 14.1176 Seconds
(without alignments)
91.421 Million cell updates/sec

Title: US-09-171-432a-48
Perfect score: 129
Sequence: 1 WLNPKINLADRLGLSGVQEIKEQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	2227	3 US-08-475-886-2	Sequence 2, Appli
2	129	100.0	2227	3 US-08-475-886-4	Sequence 4, Appli
3	129	100.0	2227	3 US-08-475-886-6	Sequence 6, Appli
4	129	100.0	2227	3 US-08-397-232-2	Sequence 2, Appli
5	129	100.0	2227	3 US-08-397-232-4	Sequence 4, Appli
6	129	100.0	2227	3 US-09-171-387-2	Sequence 2, Appli
7	129	100.0	2227	4 US-09-653-499-2	Sequence 2, Appli
8	129	100.0	2227	4 US-09-653-499-4	Sequence 4, Appli
9	129	100.0	2227	4 US-09-653-499-6	Sequence 6, Appli
10	129	100.0	2227	4 US-10-104-966-12	Sequence 12, Appli
11	129	100.0	2227	4 US-10-135-988-2	Sequence 2, Appli
12	129	100.0	2227	4 US-10-135-988-4	Sequence 4, Appli
13	129	100.0	2227	4 US-10-135-988-6	Sequence 6, Appli
14	49	38.0	2229	3 US-09-548-528-17	Sequence 17, Appli
15	49	38.0	2229	3 US-09-549-108-17	Sequence 17, Appli
16	49	38.0	2229	3 US-09-549-111-17	Sequence 17, Appli
17	49	38.0	2229	3 US-09-549-106-17	Sequence 17, Appli
18	49	38.0	2229	3 US-09-550-394-17	Sequence 17, Appli
19	47.5	36.8	699	4 US-09-252-991A-17077	Sequence 5, Appli
20	46	35.7	877	4 US-09-165-396-5	Sequence 4798, Ap
21	45	34.9	373	4 US-09-134-001C-4798	Sequence 5523, Ap
22	45	34.9	434	4 US-09-134-000C-5523	Sequence 10363, A
23	44	34.1	456	4 US-09-489-039A-10363	Sequence 6687, Ap
24	44	34.1	495	4 US-09-107-532A-6687	Sequence 2282, Ap
25	43	33.3	336	4 US-09-540-236-2282	Sequence 565, App
26	42.5	32.9	393	4 US-09-198-452A-565	Sequence 5806, Ap
27	42	32.6	179	4 US-09-134-000C-5806	Sequence 5806, Ap

ALIGNMENTS

RESULT 1

US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 WLNPKINLADRLGLSGVQEIKEQ 25
Db 956 WLNPKINLADRLGLSGVQEIKEQ 980

RESULT 2

US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232

Sequence 5935, Ap
Sequence 12, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 10946, A
Sequence 21522, A
Sequence 26240, A
Sequence 13631, A
Sequence 5, Appli
Sequence 5, Appli
Sequence 7047, Ap
Sequence 95, Appli
Sequence 9, Appli
Sequence 1, Appli
Sequence 14195, A
Sequence 1606, Ap
Sequence 5, Appli

```
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. NO. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRLMLGLSGVQEIKEQ 980

RESULT 5
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 2227
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. NO. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRLMLGLSGVQEIKEQ 980

RESULT 6
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.;
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION NUMBER: US/09/171,387
; APPLICATION NUMBER: 24-Mar-1999
```

```
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. NO. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRLMLGLSGVQEIKEQ 980

RESULT 3
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-08-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. NO. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRLMLGLSGVQEIKEQ 980

RESULT 4
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 8
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 9
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 10
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
```

```
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSCI
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/223,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 11
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 12
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
US-10-135-988-4

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 13
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 14
US-09-248-528-17
; Sequence 17, Application US/09248528C
; Patent No. 6153415
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J.
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-401
; CURRENT APPLICATION NUMBER: US/09/248,528C
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 60/083,485
; EARLIER FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
US-09-248-528-17
```

; SEQ ID NO 17
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Bacillus smithii
US-09-248-528-17

Query Match 38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.6;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 1 WLNPKKINLADRL--GLSGVOEI 22
DB 113 WENPKLVKVVKEKLEGLSPREV 136

RESULT 15
US-09-549-108-17
; Sequence 17, Application US/09549108
; Patent No. 6214603
; GENERAL INFORMATION:
; APPLICANT: Oriel, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-486
; CURRENT APPLICATION NUMBER: US/09/549,108
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Bacillus smithii
US-09-549-108-17

Query Match 38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.6;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 1 WLNPKKINLADRL--GLSGVOEI 22
DB 113 WENPKLVKVVKEKLEGLSPREV 136

Search completed: March 15, 2004, 13:26:07
Job time : 15.1176 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:23:07 ; Search time 28.9706 Seconds
(without alignments)
182.213 Million cell updates/sec

Title: US-09-171-432A-48
Perfect score: 129
Sequence: 1 WLNPKKINLADRMGLSGVQIKQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	25	10	US-09-171-432A-48
2	129	100.0	352	14	US-10-272-459-45
3	129	100.0	980	14	US-10-272-459-41
4	129	100.0	2227	9	US-09-929-955-12
5	129	100.0	2227	13	US-10-104-966-12
6	129	100.0	2227	13	US-10-135-988-2
7	129	100.0	2227	13	US-10-135-988-4
8	129	100.0	2227	13	US-10-135-988-6
9	96	74.4	20	10	US-09-171-432A-46
10	50	38.8	751	15	US-10-369-493-9973
11	48.5	37.6	1987	14	US-10-132-382-6
12	48.5	37.6	2013	14	US-10-132-382-2
13	48.5	37.6	2014	14	US-10-132-382-8
14	48.5	37.6	2040	14	US-10-132-382-4
15	48	37.2	430	15	US-10-320-797-3201

16 47 36.4 381 15 US-10-369-493-5686 Sequence 5686, Ap
17 46 35.7 877 15 US-10-369-493-22547 Sequence 22547, A
18 45.5 35.3 647 15 US-10-369-493-10283 Sequence 10283, A
19 45 34.9 189 15 US-10-108-260A-3303 Sequence 3303, Ap
20 45 34.9 275 15 US-10-108-260A-3706 Sequence 3706, Ap
21 45 34.9 309 10 US-09-510-332-67 Sequence 67, Appl
22 45 34.9 325 9 US-09-764-864-1535 Sequence 1535, Ap
23 45 34.9 357 9 US-09-938-803-16 Sequence 16, Appl
24 45 34.9 371 9 US-09-764-864-1118 Sequence 1118, Ap
25 45 34.9 400 15 US-10-108-260A-2494 Sequence 2494, A
26 45 34.9 622 14 US-10-156-761-13093 Sequence 13093, A
27 44.5 34.5 946 15 US-10-369-493-13622 Sequence 13622, A
28 44 34.1 391 15 US-10-369-493-16710 Sequence 16710, A
29 44 34.1 639 16 US-10-389-566-1869 Sequence 1869, Ap
30 44 34.1 709 15 US-10-310-154-593 Sequence 593, Appl
31 44 34.1 1770 9 US-09-801-368-298 Sequence 298, Appl
32 43 33.3 145 15 US-10-369-493-6422 Sequence 6422, Ap
33 43 33.3 145 15 US-10-369-493-6425 Sequence 6425, Ap
34 43 33.3 227 15 US-10-264-049-2405 Sequence 2405, Ap
35 43 33.3 289 14 US-10-286-264-110 Sequence 110, Ap
36 43 33.3 296 15 US-10-369-493-20999 Sequence 20999, A
37 43 33.3 358 9 US-09-861-451A-50 Sequence 50, Appl
38 43 33.3 426 14 US-10-156-761-9490 Sequence 9490, Ap
39 43 33.3 495 9 US-09-815-242-11215 Sequence 11215, A
40 43 33.3 681 15 US-10-369-493-13394 Sequence 13394, A
41 43 33.3 1098 15 US-10-104-047-2475 Sequence 2475, Ap
42 42.5 32.9 393 15 US-10-289-762-565 Sequence 565, Ap
43 42.5 32.9 687 15 US-10-369-493-1884 Sequence 1884, Ap
44 42 32.6 166 9 US-09-815-242-4942 Sequence 4942, Ap
45 42 32.6 166 9 US-09-815-242-10531 Sequence 10531, A

ALIGNMENTS

RESULT 1
US-09-171-432A-48
; Sequence 48, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single


```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..25
; OTHER INFORMATION: /label= YK-1757
US-09-171-432A-48

Query Match      100.0%; Score 129; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.5e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 1 WLNPKKINLADRMGLSGVQEIKEQ 25

RESULT 2
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match      100.0%; Score 129; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 7.4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 328 WLNPKKINLADRMGLSGVQEIKEQ 352

RESULT 3
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41
```

```
Query Match      100.0%; Score 129; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 4
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      100.0%; Score 129; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 5
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 129; DB 13; Length 2227;
```

Best Local Similarity 100.0%; Pred. No. 6.3e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 6

US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 129; DB 13; Length 2227;

Best Local Similarity 100.0%; Pred. No. 6.3e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 7

US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 129; DB 13; Length 2227;

Best Local Similarity 100.0%; Pred. No. 6.3e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 8

US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 129; DB 13; Length 2227;

Best Local Similarity 100.0%; Pred. No. 6.3e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 9

US-09-171-432A-46
; Sequence 46, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1331
US-09-171-432A-46

Query Match      74.4%; Score 96; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 KINLADRLGLSGVQEIKEQ 25
Db      1 KINLADRLGLSGVQEIKEQ 20

RESULT 10
US-10-369-493-9973
; Sequence 9973, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9973
; LENGTH: 751
; TYPE: PRT
; ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-9973

Query Match      38.8%; Score 50; DB 15; Length 751;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      4 PKKINLADRLGLSGVQEIKE 23
Db      311 PELILDDQMTGLSGVETVR 330

RESULT 11
US-10-132-382-6
; Sequence 6, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1987
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-6

Query Match      37.6%; Score 48.5; DB 14; Length 1987;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY      2 LNPKKINLADRLGLSGVQEIKE 24
Db      275 INLRKLNLSNDHNLGELPGVQSSDE 298

RESULT 12
US-10-132-382-2
; Sequence 2, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-2

Query Match      37.6%; Score 48.5; DB 14; Length 2013;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY      2 LNPKKINLADRLGLSGVQEIKE 24
Db      301 INLRKLNLSNDHNLGELPGVQSSDE 324

RESULT 13
US-10-132-382-8
; Sequence 8, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2014
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-8

Query Match      37.6%; Score 48.5; DB 14; Length 2014;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY      2 LNPKKINLADRLGLSGVQEIKE 24
Db      302 INLRKLNLSNDHNLGELPGVQSSDE 325

RESULT 14
US-10-132-382-4
; Sequence 4, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
```

```
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2040
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-4

Query Match      37.6%; Score 48.5; DB 14; Length 2040;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 LNPXKINLADRMGLG-LSGVQEIKE 24
   |||||:|||||
Db 328 INLRKLNLSDNHGLGELPGVQSSDE 351

RESULT 15
US-10-320-797-3201
; Sequence 3201, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; FILE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3201
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3201

Query Match      37.2%; Score 48; DB 15; Length 430;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLGSG 18
   |||||:|||||
Db 302 WPNLKHNLSDCLLGSG 319

Search completed: March 15, 2004, 13:53:28
Job time : 28.9706 secs
```

Sequence 48, Appl
Sequence 48, Appl
Sequence 45, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 46, Appl
Sequence 46, Appl
Sequence 175643,
Sequence 9973, Ap
Sequence 9973, Ap
Sequence 16538, A
Sequence 4455, Ap
Sequence 146066,
Sequence 13902, A
Sequence 6, Appl
Sequence 2, Appl
Sequence 8, Appl
Sequence 18, Appl
Sequence 8, Appl
Sequence 4, Appl
Sequence 4704, Ap
Sequence 4704, Ap
Sequence 4821, Ap
Sequence 4722, Ap
Sequence 139229,
Sequence 3201, Ap
Sequence 3201, Ap
Sequence 3201, Ap
Sequence 67150, A
Sequence 51412, A
Sequence 48239, A
Sequence 5321, Ap
Sequence 17077, A
Sequence 17077, A
Sequence 138302,
Sequence 138307,
Sequence 7, Appl

1 129 100.0 25 15 US-09-171-432A-48
2 129 100.0 25 26 US-10-031-088A-48
3 129 100.0 352 1 PCT-US02-33121-45
4 129 100.0 352 28 US-10-272-459-45
5 129 100.0 980 1 PCT-US02-33121-41
6 129 100.0 980 28 US-10-272-459-41
7 129 100.0 2227 7 US-08-397-232-2
8 129 100.0 2227 8 US-08-475-886-2
9 129 100.0 2227 8 US-08-475-886-4
10 129 100.0 2227 21 US-09-705-547-12
11 129 100.0 2227 24 US-09-929-955-12
12 129 100.0 2227 32 US-10-719-619-12
13 129 100.0 2227 32 US-60-229-175-12
14 96 74.4 20 15 US-09-171-432A-46
15 96 74.4 20 26 US-10-031-088A-46
16 50 38.8 253 30 US-10-424-599-175643
17 50 38.8 751 29 US-10-369-493-9973
18 50 38.8 751 33 US-60-360-039-9973
19 50 38.8 1176 31 US-10-679-063-16538
20 49 38.0 327 22 US-09-791-537-146066
21 49 38.0 474 22 US-09-791-537-146066
22 49 38.0 579 31 US-10-679-063-13902
23 48.5 37.6 1987 27 US-10-132-382-6
24 48.5 37.6 2013 27 US-10-132-382-2
25 48.5 37.6 2014 27 US-10-132-382-8
26 48.5 37.6 2014 28 US-10-220-955-18
27 48.5 37.6 2014 30 US-10-415-011-8
28 48.5 37.6 2040 27 US-10-132-382-4
29 48 37.2 213 20 US-09-614-150-4704
30 48 37.2 213 20 US-09-614-150A-4704
31 48 37.2 213 33 US-60-167-217-4821
32 48 37.2 213 33 US-60-191-637-4722
33 48 37.2 334 22 US-09-791-537-139229
34 48 37.2 430 1 PCT-US02-40225-3201
35 48 37.2 430 29 US-10-320-797-3201
36 48 37.2 430 33 US-60-341-261-3201
37 48 37.2 622 22 US-09-791-537-67150
38 48 37.2 841 28 US-10-282-122A-51412
39 47.5 36.8 603 28 US-10-282-122A-48239
40 47.5 36.8 640 19 US-09-540-203B-5321
41 47.5 36.8 699 29 US-10-366-683-17077
42 47.5 36.8 699 30 US-10-419-128-17077
43 47.5 36.8 981 30 US-10-437-963-138302
44 47.5 36.8 1030 30 US-10-437-963-138307
45 47 36.4 145 31 US-10-657-740-7

ALIGNMENTS

RESULT 1
US-09-171-432A-48
; Sequence 48, Application US/09171432A
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khuyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Klpatrik Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:16:32 ; Search time 200.735 Seconds
(without alignments)
121.560 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 WLNPKINLADRLGLSGVQIEKQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/2/paa/pctus COMB.pcp.*
- 2: /cgn2_6/ptodata/2/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/2/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/2/paa/US081 COMB.pcp.*
- 6: /cgn2_6/ptodata/2/paa/US082 COMB.pcp.*
- 7: /cgn2_6/ptodata/2/paa/US083 COMB.pcp.*
- 8: /cgn2_6/ptodata/2/paa/US084 COMB.pcp.*
- 9: /cgn2_6/ptodata/2/paa/US085 COMB.pcp.*
- 10: /cgn2_6/ptodata/2/paa/US086 COMB.pcp.*
- 11: /cgn2_6/ptodata/2/paa/US087 COMB.pcp.*
- 12: /cgn2_6/ptodata/2/paa/US088 COMB.pcp.*
- 13: /cgn2_6/ptodata/2/paa/US089 COMB.pcp.*
- 14: /cgn2_6/ptodata/2/paa/US090 COMB.pcp.*
- 15: /cgn2_6/ptodata/2/paa/US091 COMB.pcp.*
- 16: /cgn2_6/ptodata/2/paa/US092 COMB.pcp.*
- 17: /cgn2_6/ptodata/2/paa/US093 COMB.pcp.*
- 18: /cgn2_6/ptodata/2/paa/US094 COMB.pcp.*
- 19: /cgn2_6/ptodata/2/paa/US095 COMB.pcp.*
- 20: /cgn2_6/ptodata/2/paa/US096 COMB.pcp.*
- 21: /cgn2_6/ptodata/2/paa/US097A COMB.pcp.*
- 22: /cgn2_6/ptodata/2/paa/US097B COMB.pcp.*
- 23: /cgn2_6/ptodata/2/paa/US098 COMB.pcp.*
- 24: /cgn2_6/ptodata/2/paa/US099A COMB.pcp.*
- 25: /cgn2_6/ptodata/2/paa/US099B COMB.pcp.*
- 26: /cgn2_6/ptodata/2/paa/US100 COMB.pcp.*
- 27: /cgn2_6/ptodata/2/paa/US101 COMB.pcp.*
- 28: /cgn2_6/ptodata/2/paa/US102 COMB.pcp.*
- 29: /cgn2_6/ptodata/2/paa/US103 COMB.pcp.*
- 30: /cgn2_6/ptodata/2/paa/US104 COMB.pcp.*
- 31: /cgn2_6/ptodata/2/paa/US106 COMB.pcp.*
- 32: /cgn2_6/ptodata/2/paa/US107 COMB.pcp.*
- 33: /cgn2_6/ptodata/2/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	%	Score	Match	Length	ID	Description
--------	-------	---	-------	-------	--------	----	-------------

```
/ FILING DATE: 23-NOV-1998
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Greene, Jamie L.
/ REGISTRATION NUMBER: 32,467
/ REFERENCE/DOCKET NUMBER: 03063-02311US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (404) 949-2400
/ TELEFAX: (404) 949-2499
/ INFORMATION FOR SEQ ID NO: 48:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 25 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..25
/ OTHER INFORMATION: /label=YK-1757
/
US-09-171-432A-48

Query Match 100.0%; Score 129; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 9.7e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
DB 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25

RESULT 2
US-10-031-088A-48
/ Sequence 48, Application US/10031088A
/ GENERAL INFORMATION:
/ APPLICANT: Howard A. Fields
/ TITLE OF INVENTION: Synthetic Peptides Immunoreactive With Hepatitis A
/ FILE REFERENCE: 14114.0342P1
/ CURRENT FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: US/10/031,088A
/ PRIOR FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: PCT/US00/19267
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/144,412
/ PRIOR FILING DATE: 1999-07-15
/ NUMBER OF SEQ ID NOS: 94
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 48
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic HAV
/ OTHER INFORMATION: P2A Peptide
/
US-10-031-088A-48

Query Match 100.0%; Score 129; DB 26; Length 25;
Best Local Similarity 100.0%; Pred. No. 9.7e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
DB 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25

RESULT 3
PCT-US02-33121-45
/ Sequence 45, Application PC/TUS0233121
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON CORPORATION et al.
/ TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
/ FILE REFERENCE: P17955.003 / 2301-17955.40
/ CURRENT APPLICATION NUMBER: PCT/US02/33121
/ CURRENT FILING DATE: 2002-10-15
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 41
/ LENGTH: 980
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
/ OTHER INFORMATION: of 38.8 kDa
/
PCT-US02-33121-45

Query Match 100.0%; Score 129; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
DB 328 WLNPKKINLADRLMLGSLGVQEIKEQ 352

RESULT 4
US-10-272-459-45
/ Sequence 45, Application US/10272459
/ GENERAL INFORMATION:
/ APPLICANT: PICHUANES, Sergio
/ TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
/ FILE REFERENCE: P17955.002 / 2301-17955
/ CURRENT APPLICATION NUMBER: US/10/272,459
/ CURRENT FILING DATE: 2002-10-15
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 45
/ LENGTH: 352
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
/ OTHER INFORMATION: of 38.8 kDa
/
US-10-272-459-45

Query Match 100.0%; Score 129; DB 28; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
DB 328 WLNPKKINLADRLMLGSLGVQEIKEQ 352

RESULT 5
PCT-US02-33121-41
/ Sequence 41, Application PC/TUS0233121
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON CORPORATION et al.
/ TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
/ FILE REFERENCE: P17955.003 / 2301-17955.40
/ CURRENT APPLICATION NUMBER: PCT/US02/33121
/ CURRENT FILING DATE: 2002-10-15
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 41
/ LENGTH: 980
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
/ OTHER INFORMATION: of 38.8 kDa
/
US-10-272-459-45

Query Match 100.0%; Score 129; DB 28; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
DB 328 WLNPKKINLADRLMLGSLGVQEIKEQ 352
```

OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
OTHER INFORMATION: of 115.5 kDa
PCT-US02-33121-41

Query Match 100.0%; Score 129; DB 1; Length 980;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 6
US-10-272-459-41
Sequence 41, Application US/10272459
GENERAL INFORMATION:
APPLICANT: PICHUANES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
FILE REFERENCE: PP17955.002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 41
LENGTH: 980
TYPE: PRI
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
OTHER INFORMATION: of 115.5 kDa
US-10-272-459-41

Query Match 100.0%; Score 129; DB 28; Length 980;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 7
US-08-397-232-2
Sequence 2, Application US/08397232
GENERAL INFORMATION:
APPLICANT: Funkhouser, Ann. W.
APPLICANT: Emerson, Suzanne U.
APPLICANT: Purcell, Robert H.
APPLICANT: D'Hondt, Eric
TITLE OF INVENTION: Hepatitis A Virus Vaccines
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: PO Box 457 Spring House Corporate Cntr
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/397,232
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/947,338
FILING DATE: 18-SEP-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US93/08610
FILING DATE: 17-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: SBPC50110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-397-232-2

Query Match 100.0%; Score 129; DB 7; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 8
US-08-475-886-2
Sequence 2, Application US/08475886
GENERAL INFORMATION:
APPLICANT: Funkhouser, Ann W.
APPLICANT: Emerson, Suzanne U.
APPLICANT: Purcell, Robert H.
APPLICANT: D'Hondt, Eric
TITLE OF INVENTION: Hepatitis A Virus Vaccines
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: PO Box 457 Spring House Corporate Cntr
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,886
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/947,338
FILING DATE: 18-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/397,232
FILING DATE: 10-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: SBPC50110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-475-886-2

Query Match 100.0%; Score 129; DB 8; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
|||||
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980
|||||

RESULT 9
US-08-475-886-4
; Sequence 4, Application US/08475886
; GENERAL INFORMATION:
; APPLICANT: Funkhouser, Ann W.
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; APPLICANT: D'Hondt, Eric
; TITLE OF INVENTION: Hepatitis A Virus Vaccines
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: PO Box 457 Spring House Corporate Cntr
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,886
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/947,338
; FILING DATE: 18-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/397,232
; FILING DATE: 10-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: SBEP50110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-475-886-4

Query Match 100.0%; Score 129; DB 8; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
|||||
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980
|||||

RESULT 10
US-09-705-547-12
; Sequence 12, Application US/09705547
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; METHODS OF USE THEREOF

FILE REFERENCE: TRIPEP.023AUS
; CURRENT APPLICATION NUMBER: US/09/705,547
; CURRENT FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-705-547-12

Query Match 100.0%; Score 129; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
|||||
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980
|||||

RESULT 11
US-09-929-955-12
; Sequence 12, Application US/09929955
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 129; DB 24; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
|||||
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980
|||||

RESULT 12
US-10-719-619-12
; Sequence 12, Application US/10719619
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1C
; CURRENT APPLICATION NUMBER: US/10/719,619
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 10/104,966
; PRIOR FILING DATE: 2002-03-22


```
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-719-619-12

Query Match          100.0%; Score 129; DB 32; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
Db 956 WLNPKKINLADRLMLGSLGVQEIKEQ 980

RESULT 13
US-60-229-175-12
; Sequence 12, Application US/60229175
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.023PR
; CURRENT APPLICATION NUMBER: US/60/229,175
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-60-229-175-12

Query Match          100.0%; Score 129; DB 33; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGSLGVQEIKEQ 25
Db 956 WLNPKKINLADRLMLGSLGVQEIKEQ 980

RESULT 14
US-09-171-432A-46
; Sequence 46, Application US/09171432A
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1331
US-09-171-432A-46

Query Match          74.4%; Score 96; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 KINLADRLMLGSLGVQEIKEQ 25
Db 1 KINLADRLMLGSLGVQEIKEQ 20

RESULT 15
US-10-031-088A-46
; Sequence 46, Application US/10031088A
; GENERAL INFORMATION:
; APPLICANT: Howard A. Fields
; APPLICANT: Yury E. Khudyakov
; TITLE OF INVENTION: Synthetic Peptides Immunoreactive With Hepatitis A
; TITLE OF INVENTION: Virus Antibodies
; FILE REFERENCE: 14114.0342P1
; CURRENT APPLICATION NUMBER: US/10/031,088A
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: PCT/US00/19267
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/144,412
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic HAV
; OTHER INFORMATION: P2A Peptide
US-10-031-088A-46

Query Match          74.4%; Score 96; DB 26; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 KINLADRLMLGSLGVQEIKEQ 25
Db 1 KINLADRLMLGSLGVQEIKEQ 20

Search completed: March 15, 2004, 13:49:07
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:22:37 ; Search time 6.47059 Seconds
(without alignments)
55.073 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 130349 segs, 14254074 residues

Total number of hits satisfying chosen parameters: 130349

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:**

- 1: /cgn2_6/ptodata/1/paa/US05_NEW_COMB.pep:**
- 2: /cgn2_6/ptodata/1/paa/US05_NEW_COMB.pep:**
- 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:**
- 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:**
- 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:**
- 6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:**
- 7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	44.2	522	6	US-10-767-701-46006
2	44	34.1	62	6	US-10-767-701-47926
3	44	34.1	495	6	US-10-417-884A-6687
4	43	33.3	151	6	US-10-767-701-59042
5	43	33.3	622	6	US-10-767-701-45235
6	42	32.6	183	6	US-10-417-884A-5935
7	41	31.8	111	6	US-10-767-701-44063
8	41	31.8	156	6	US-10-767-701-40341
9	41	31.8	162	6	US-10-767-701-59711
10	41	31.8	305	6	US-10-363-552-65
11	41	31.8	514	6	US-10-778-469-3
12	41	31.8	537	6	US-10-767-471-1304
13	41	31.8	537	6	US-10-767-471-1306
14	41	31.8	537	6	US-10-767-471-1308
15	41	31.8	641	6	US-10-767-471-1303
16	41	31.8	641	6	US-10-767-471-1307
17	41	31.8	687	6	US-10-767-471-1302
18	41	31.8	687	6	US-10-767-471-1305
19	41	31.8	708	6	US-10-775-972-369
20	40.5	31.4	55	6	US-10-767-701-50935
21	40.5	31.4	134	6	US-10-767-701-43829
22	40.5	31.4	170	6	US-10-767-701-59036
23	40.5	31.4	367	6	US-10-786-720-48
24	40	31.0	117	6	US-10-767-701-44815
25	40	31.0	144	6	US-10-417-884A-7063
26	40	31.0	181	6	US-10-767-701-37982

Query Match 44.2%; Score 57; DB 6; Length 522;
Best Local Similarity 40.0%; Pred. No. 0.27;
Matches 12; Conservative 6; Mismatches 6; Indels 6; Gaps 1;

Qy 1 WLNPKKINLADRMGLSGVQ-----RIKE 24
Db 39 WLRPRRLNLRALRAQGLSGTGYLFPAGDLKE 68

RESULT 2

US-10-767-701-47926
; Sequence 47926, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 47926

; LENGTH: 62

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3480-036-P1-K1-D11.pep

ALIGNMENTS

RESULT 1

US-10-767-701-46006

; Sequence 46006, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 46006

; LENGTH: 522

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2930_1.pep

Sequence 44360, A
Sequence 37196, A
Sequence 43553, A
Sequence 168, App
Sequence 7, Appl
Sequence 4552, Ap
Sequence 23, Appl
Sequence 8370, Ap
Sequence 10743, A
Sequence 54370, A
Sequence 59189, A
Sequence 32852, A
Sequence 6064, Ap
Sequence 36537, A
Sequence 25, Appl
Sequence 4438, Ap
Sequence 4439, Ap
Sequence 178, App

27 40 31.0 239 6 US-10-767-701-44360
28 40 31.0 254 6 US-10-767-701-37198
29 40 31.0 422 6 US-10-767-701-43553
30 40 31.0 472 1 PCT-US03-13562-168
31 40 31.0 919 6 US-10-398-037-7
32 40 31.0 1297 6 US-10-417-884A-4552
33 40 31.0 1673 5 US-09-889-874A-23
34 39 30.2 99 6 US-10-100-683-8370
35 39 30.2 99 6 US-10-100-683-10743
36 39 30.2 170 6 US-10-767-701-54370
37 39 30.2 174 6 US-10-767-701-59189
38 39 30.2 217 6 US-10-767-701-32852
39 39 30.2 229 6 US-10-417-884A-6064
40 39 30.2 251 6 US-10-767-701-36537
41 39 30.2 280 5 US-09-687-864A-25
42 39 30.2 294 5 US-09-687-864A-24
43 39 30.2 311 6 US-10-417-884A-4438
44 39 30.2 311 6 US-10-417-884A-4439
45 39 30.2 327 1 PCT-US04-05654-178

```
US-10-767-701-47926
Query Match      34.1%; Score 44; DB 6; Length 62;
Best Local Similarity 36.4%; Pred. No. 2.4;
Matches 8; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 4 PKKINLADRMGLSGVQBIKEQ 25
Db 5 PPGLNLADCGVGAAGLEQLIQ 26

RESULT 3
US-10-417-884A-6687
; Sequence 6687, Application US/10417884A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/417,884A
; FILING DATE: 17-Apr-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6687:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...495
; SEQUENCE DESCRIPTION: SEQ ID NO: 6687:
US-10-417-884A-6687

Query Match      34.1%; Score 44; DB 6; Length 495;
Best Local Similarity 27.3%; Pred. No. 29;
Matches 9; Conservative 9; Mismatches 7; Indels 8; Gaps 1;

QY 1 WLNPKINLADRMGLSG-----VQBIKEQ 25
Db 21 WMEMDRLYQNRLMGIGDALESVEVRPKKE 53

RESULT 4
US-10-767-701-59042
; Sequence 59042, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 59042
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; NAME/KEY: unsure
; LOCATION: (1)..(151)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: 6859111.pep
US-10-767-701-59042

Query Match      33.3%; Score 43; DB 6; Length 151;
Best Local Similarity 44.4%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 WLNPKINLADRMGLSG 18
Db 40 WLRPRRLGRALRSQGLPG 57

RESULT 5
US-10-767-701-45235
; Sequence 45235, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45235
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(622)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C14268_1.pep
US-10-767-701-45235

Query Match      33.3%; Score 43; DB 6; Length 622;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 9; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 4 PKKINLADRMGLSG 18
Db 529 PKKVNLD--IGIV 541

RESULT 6
US-10-417-884A-5935
; Sequence 5935, Application US/10417884A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
```

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29

PRIOR APPLICATION DATA:
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5935:
SEQUENCE CHARACTERISTICS:
LENGTH: 183 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...183
SEQUENCE DESCRIPTION: SEQ ID NO: 5935:
US-10-417-884A-5935

Query Match 32.6%; Score 42; DB 6; Length 183;
Best Local Similarity 27.3%; Pred. No. 18;
Matches 6; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 WLNPKINLADRMGLSGVQEI 22
DB 20 YIDPKHLEEDRVAINRVTKV 41

RESULT 7
US-10-767-701-44063
; Sequence 44063, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 44063
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:

US-09-1/1-4328-10.1a.pn

OTHER INFORMATION: Clone ID: SORBI-28MAY03-C63072_1.ppe
US-10-767-701-44063

Query Match 31.8%; Score 41; DB 6; Length 111;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 PKKINLADRMGL 15
DB 91 PKRCNWADRPVG 102

RESULT 8
US-10-767-701-40341
; Sequence 40341, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40341
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; NAME/KEY: unsure
; LOCATION: (1)...(156)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C56808_1.ppe
US-10-767-701-40341

Query Match 31.8%; Score 41; DB 6; Length 156;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 LADRMGLSGVQEI 22
DB 112 MADHLGCSGAEDI 125

RESULT 9
US-10-767-701-59711
; Sequence 59711, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 59711
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 7535660.ppe
US-10-767-701-59711

Query Match 31.8%; Score 41; DB 6; Length 162;
Best Local Similarity 52.2%; Pred. No. 23;
Matches 12; Conservative 1; Mismatches 4; Indels 6; Gaps 1;

QY 2 LNPKNLADRMGLSGVQEI 24
DB 112 MADHLGCSGAEDI 125

Db 44 LNPCKSD-----GLSGKVEKE 60

RESULT 10

US-10-363-552-65

Sequence 65, Application US/10363552

GENERAL INFORMATION:

APPLICANT: STUMPP, MICHAEL TOBIAS

TITLE OF INVENTION: COLLECTIONS OF REPEAT PROTEINS COMPRISING REPEAT

TITLE OF INVENTION: MODULES

FILE REFERENCE: 37629-0075

CURRENT FILING DATE: 2003-11-24

PRIOR APPLICATION NUMBER: PCT/EP01/10454

PRIOR FILING DATE: 2001-09-10

NUMBER OF SEQ ID NOS: 77

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 65

LENGTH: 305

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Translated amino acid

OTHER INFORMATION: sequence of the synthetic NcoI-HindIII insert

US-10-363-552-65

Query Match 31.8%; Score 41; DB 6; Length 305;

Best Local Similarity 45.0%; Pred. No. 48;

Matches 9; Conservative 6; Mismatches 3; Indels 2; Gaps 1;

QY 3 NP--KKINLADRMGLSGVQ 20

Db 127 NPSRELNSDNKLGDAVR 146

RESULT 11

US-10-778-469-3

Sequence 3, Application US/10778469

GENERAL INFORMATION:

APPLICANT: Bisgard-Frantzen, Henrik

Svendsen, Allan

Borchert, Torben Vedel

TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Novo Nordisk of North America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/778,469

FILING DATE: 12-Feb-2004

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/354,191A

FILING DATE: 15-JULY-1999

APPLICATION NUMBER: 08/600,656

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4318.204-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 514 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-778-469-3

Query Match 31.8%; Score 41; DB 6; Length 514;

Best Local Similarity 30.4%; Pred. No. 91;

Matches 7; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEI 23

Db 115 WVDAVEVNPDSRQNEISGTQIQ 137

RESULT 12

US-10-767-471-1304

Sequence 1304, Application US/10767471

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001505

CURRENT APPLICATION NUMBER: US/10/767,471

CURRENT FILING DATE: 2004-01-30

NUMBER OF SEQ ID NOS: 50231

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1304

LENGTH: 537

TYPE: PRT

ORGANISM: Homo sapiens

US-10-767-471-1304

Query Match 31.8%; Score 41; DB 6; Length 537;

Best Local Similarity 53.3%; Pred. No. 96;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 PKKINLADRMGLSG 18

Db 270 PRKANLDSMEGSPG 284

RESULT 13

US-10-767-471-1306

Sequence 1306, Application US/10767471

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001505

CURRENT APPLICATION NUMBER: US/10/767,471

CURRENT FILING DATE: 2004-01-30

NUMBER OF SEQ ID NOS: 50231

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1306

LENGTH: 537

TYPE: PRT

ORGANISM: Homo sapiens

US-10-767-471-1306

Query Match 31.8%; Score 41; DB 6; Length 537;

Best Local Similarity 53.3%; Pred. No. 96;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 PKKINLADRMGLSG 18

Db 270 PRKANLDSMEGSPG 284

RESULT 14

Search completed: March 15, 2004, 13:50:00
Job time : 7.47059 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:50:08 ; Search time 12.5 Seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-09-171-432A-48
Perfect score: 129
Sequence: 1 WLNPKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: piri:*

2: piri:*

3: piri:*

4: piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	2227	1 GNNYHM	genome polyprotein
2	129	100.0	2227	1 GNNYHR	genome polyprotein
3	129	100.0	2227	1 GNNYMK	genome polyprotein
4	129	100.0	2227	1 GNNYHB	genome polyprotein
5	129	100.0	2230	1 GNNYSA	genome polyprotein
6	51.5	39.9	656	2 B82056	glutathione-regula
7	50	38.8	172	2 S33416	heat shock protein
8	49	38.0	327	1 RDVZAS	ribonucleoside-dip
9	48	37.2	158	2 S64321	hypothetical prote
10	47.5	36.8	482	2 S27608	cysteine proteinas
11	47.5	36.8	653	2 G82971	probable ferredoxi
12	47	36.4	145	2 B25199	heat shock 16k pro
13	47	36.4	381	2 T18876	hypothetical prote
14	46.5	36.0	146	2 B70142	ribosomal protein
15	46	35.7	342	2 H81317	probable lipopolys
16	46	35.7	342	2 H86336	hypothetical prote
17	46	35.7	518	2 S73432	MG096 homolog D09
18	46	35.7	632	2 S73431	MG288 homolog D09
19	46	35.7	671	2 AF0042	probable oxidoredu
20	46	35.7	877	2 S58824	probable membrane
21	45.5	35.3	201	2 H81274	amidotransferase C
22	45.5	35.3	1027	2 T43024	ceOB protein - Bur
23	45	34.9	1042	2 H70203	isoleucine-tRNA li
24	45	34.9	1061	2 H90084	hypothetical prote
25	44	34.1	152	2 S07505	endodeoxyribonucle
26	44	34.1	165	2 S73194	hypothetical prote
27	44	34.1	278	2 B82388	probable maltose o
28	44	34.1	370	2 H70423	oxygen-independent
29	44	34.1	408	2 AC2295	succinyl-CoA synth

30 44 34.1 412 2 C96816
31 44 34.1 414 2 T06303
32 44 34.1 512 2 T48462
33 44 34.1 646 2 S72609
34 44 34.1 653 2 G70683
35 44 34.1 677 2 B82870
36 44 34.1 788 2 S67595
37 44 34.1 994 2 S18739
38 44 34.1 1770 2 S56221
39 43.5 33.7 388 1 S72995
40 43.5 33.7 395 2 AD0380
41 43.5 33.7 482 2 A38533
42 43 33.3 145 2 B24289
43 43 33.3 148 2 G69125
44 43 33.3 180 2 AC0056
45 43 33.3 207 2 C49905

ALIGNMENTS

RESULT 1

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
E; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
A:Accession: A25981
R:Conen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with d
A:Reference number: A25981; MUID:87061253; PMID:3023706
A:Accession: A25981
A:Molecule type: Genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:MI4707; NID:G329582; PIDN:AAA45465.1; PID:G329583
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltran
E:1-23/Product: coat protein 1A #status predicted <VP4>
E:24-245/Product: coat protein 1B #status predicted <VP2>
E:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP1>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKINLADRMGLSGVQEIKEQ 25

Db 956 WLNPKINLADRMGLSGVQEIKEQ 980

RESULT 2

GNNYHR

genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pr
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
A:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nesl
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.


```
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSQAKISLFYTEHEIMKFS 20
Db      834 FSQAKISLFYTEHEIMKFS 853

RESULT 15
US-09-489-039A-8111
; Sequence 8111, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8111
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8111

Query Match      45.5%; Score 46; DB 4; Length 741;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 FSQAKISLFYTEHEIMK 18
Db      345 FREAEISQLYTKEHPTYK 362

Search completed: March 15, 2004, 13:26:04
Job time : 11.2941 secs
```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1319
US-09-171-432A-43

Query Match      100.0%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSOAKISLFYTEHEIMKFS 20
Db      1 FSOAKISLFYTEHEIMKFS 20

RESULT 2
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
US-10-272-459-45

Query Match      100.0%; Score 101; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSOAKISLFYTEHEIMKFS 20
Db      206 FSOAKISLFYTEHEIMKFS 225

RESULT 3
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 115.5 kDa
US-10-272-459-41

Query Match      100.0%; Score 101; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSOAKISLFYTEHEIMKFS 20
Db      834 FSOAKISLFYTEHEIMKFS 853

RESULT 4
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSOAKISLFYTEHEIMKFS 20
Db      834 FSOAKISLFYTEHEIMKFS 853

RESULT 5
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 101; DB 13; Length 2227;
```

```
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEHEIMKFS 20
|||||
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 6
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEHEIMKFS 20
|||||
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 7
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEHEIMKFS 20
|||||
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 8
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEHEIMKFS 20
|||||
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 9
US-09-252-088-6
; Sequence 6, Application US/09252088
; Publication No. US20030031682A1
; GENERAL INFORMATION:
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: RIOUX, Clment
; APPLICANT: BOYER, Martine
; APPLICANT: CHARLEBOIS, Isabelle
; APPLICANT: HAMEL, Jose
; APPLICANT: MARTIN, Denis
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 8331-9002
; CURRENT APPLICATION NUMBER: US/09/252,088
; CURRENT FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: US/60/075,425
; EARLIER FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 518
; TYPE: PRT
; ORGANISM: group B streptococcus
US-09-252-088-6

Query Match 46.5%; Score 47; DB 10; Length 518;
Best Local Similarity 45.0%; Pred. No. 20;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEHEIMKFS 20
|||||
Db 371 FQAPIALVKNABEVVAF 390

RESULT 10
US-10-340-792-6
```

```
; Sequence 6, Application US/10340792
; Publication No. US20030228323A1
; GENERAL INFORMATION:
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: RIOUX, Clement
; APPLICANT: BOYER, Martine
; APPLICANT: CHARLEBOIS, Isabelle
; APPLICANT: HAMEL, Josee
; APPLICANT: MARTIN, Denis
; APPLICANT: HAMMILL, Joseph
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
; CURRENT APPLICATION NUMBER: US/10/340,792
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US/09/252,088B
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/60/075,425
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 518
; TYPE: PRT
; ORGANISM: group B streptococcus
; US-10-340-792-6

Query Match 46.5%; Score 47; DB 15; Length 518;
Best Local Similarity 45.0%; Pred. No. 20;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FQAKISLFYTEHEIMKFS 20
DB 371 FQAPIALVKNAHEVVAFA 390

RESULT 11
US-09-252-088-5
; Sequence 5, Application US/09252088
; Publication No. US20030031682A1
; GENERAL INFORMATION:
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: RIOUX, Clement
; APPLICANT: BOYER, Martine
; APPLICANT: CHARLEBOIS, Isabelle
; APPLICANT: HAMEL, Josee
; APPLICANT: MARTIN, Denis
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
; CURRENT APPLICATION NUMBER: US/09/252,088
; CURRENT FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: US/60/075,425
; EARLIER FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 816
; TYPE: PRT
; ORGANISM: group B streptococcus
; US-09-252-088-5

Query Match 46.5%; Score 47; DB 10; Length 816;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FQAKISLFYTEHEIMKFS 20
DB 669 FQAPIALVKNAHEVVAFA 688

RESULT 12
US-10-340-792-5
; Sequence 5, Application US/10340792
; Publication No. US20030228323A1
; GENERAL INFORMATION:
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: RIOUX, Clement
; APPLICANT: BOYER, Martine
; APPLICANT: CHARLEBOIS, Isabelle
; APPLICANT: HAMEL, Josee
; APPLICANT: MARTIN, Denis
; APPLICANT: HAMMILL, Joseph
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
; CURRENT APPLICATION NUMBER: US/10/340,792
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US/09/252,088B
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/60/075,425
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 816
; TYPE: PRT
; ORGANISM: group B streptococcus
; US-10-340-792-5

Query Match 46.5%; Score 47; DB 15; Length 816;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FQAKISLFYTEHEIMKFS 20
DB 669 FQAPIALVKNAHEVVAFA 688

RESULT 13
US-09-171-432A-42
; Sequence 42, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171.432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1318
US-09-171-432A-42

Query Match      41.6%; Score 42; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 FSQAKISLF 9
   |||||
Db 12 FSQAKISLF 20

RESULT 14
US-10-369-493-21606
; Sequence 21606, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21606
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Pyrococcus abyssi
US-10-369-493-21606

Query Match      40.6%; Score 41; DB 15; Length 144;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QAKISLFYTEEHE 15
   :|||:
Db 18 KVKEGLYTKHE 30

RESULT 15
US-10-369-493-3295
; Sequence 3295, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3295
; LENGTH: 1260
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
```

```
; LOCATION: (1)..(1260)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3295

Query Match      40.6%; Score 41; DB 15; Length 1260;
Best Local Similarity 44.4%; Pred. No. 5.1e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 SQAKISLFYTEEHEIMKPF 19
   :|||:
Db 393 SERKIKLFSSOHLIQOF 410

Search completed: March 15, 2004, 13:53:27
Job time : 23.1765 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:50:08 ; Search time 10 Seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-09-171-432A-43
Perfect score: 101
Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	101	100.0	1358	2 A03905	genome polyprotein
2	101	100.0	2227	1 GNNYHM	genome polyprotein
3	101	100.0	2227	1 GNNYHR	genome polyprotein
4	101	100.0	2227	1 GNNYMK	genome polyprotein
5	101	100.0	2227	1 GNNYHB	genome polyprotein
6	97	96.0	852	1 GNNYHA	genome polyprotein
7	88	87.1	2230	1 GNNYSA	genome polyprotein
8	47	46.5	378	2 F64300	formate dehydrogen
9	43	42.6	325	2 H71869	hypothetical prote
10	43	42.6	410	2 T22413	hypothetical prote
11	42	41.6	125	2 AB1747	glycine cleavage s
12	42	41.6	125	2 A11377	glycine cleavage s
13	42	41.6	187	2 S10844	hypothetical prote
14	42	41.6	545	2 S10842	gene ND5 intron 2
15	42	41.6	1265	2 F88886	protein F52G2.2 [i
16	41.5	41.1	288	2 S64243	hypothetical prote
17	41	40.6	144	2 E75128	glycine cleavage s
18	41	40.6	190	2 C90529	nuclease, lipoprot
19	41	40.6	209	2 S65300	hypothetical prote
20	41	40.6	304	2 C72324	homoserine O-succi
21	41	40.6	465	2 B97235	deacetylase/diipep
22	41	40.6	483	2 G84113	hypothetical prote
23	41	40.6	557	2 S58164	glucose-6-phosphat
24	41	40.6	820	2 T51510	hypothetical prote
25	41	40.6	1040	1 A38306	alpha-mannosidase
26	41	40.6	2431	1 MNVVSF	nonstructural poly
27	40.5	40.1	251	2 T13104	minor tail protein
28	40.5	40.1	419	2 T25577	hypothetical prote
29	40	39.6	142	2 C70059	hypothetical prote

ALIGNMENTS

RESULT 1

A03905
genome polyprotein (version 2) - human hepatitis A virus (fragments)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein
C;Species: human hepatitis A virus
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996
C;Accession: A03905
P;Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinsto
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985
A;Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A;Reference number: A03905; MUID:85166289; PMID:2984684
A;Accession: A03905
A;Molecule type: genomic RNA
A;Residues: 1-1358 <BAR>
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein
P;1-245/Product: coat protein 1A #status predicted <C1A>
P;246-491/Product: coat protein 1B #status predicted <C1B>
P;492-836/Product: coat protein 1C #status predicted <C1C>
P;837-854/Product: core protein 2A (fragment) #status predicted <C2A>
P;855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 101; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSQAKISLFYTEHEIMKFS 20
|||||
Db 834 FSQAKISLFYTEHEIMKFS 853

RESULT 2

GNNYHM
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C;Accession: A25581
P;Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A;Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with di
A;Reference number: A25981; MUID:87061253; PMID:3023706
A;Accession: A25981
A;Molecule type: genomic RNA
A;Residues: 1-2227 <COH>
A;Cross-references: EMBL:M14707; NID:G329582; PIDN:AAA45465.1; PID:G329583
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltran
P;1-23/Product: coat protein 1A #status predicted <VP4>
P;24-245/Product: coat protein 1B #status predicted <VP2>

F;246-491/Product: coat protein 1C #status predicted <VP3>
F;492-791/Product: coat protein 1D #status predicted <VP1>
F;792-980/Product: coat protein 2A #status predicted <VP2>
F;981-1087/Product: coat protein 2B #status predicted <C2B>
F;1088-1422/Product: coat protein 2C #status predicted <C2C>
F;1423-1496/Product: protein 3A #status predicted <C3A>
F;1497-1519/Product: protein 3B #status predicted <C3B>
F;1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F;1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
|||||
Db 834 FSQAKISLFYTEHEIMKFS 853

RESULT 3
GNNVHR
genome polyprotein - human hepatitis A virus
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core pro
NA polymerase (EC 2.7.7.48); protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C;Accession: A03903
R;Ajarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nes
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A;Title: Primary structure and gene organization of human hepatitis A virus.
A;Reference number: A03903; MUID:85190549; PMID:2986127
A;Accession: A03903
A;Molecule type: genomic RNA
A;Residues: 1-2227 <NAJ>
A;Cross-references: GB:K02990; NID:G329596; PIDN:AAA45472.1; PID:G329597
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleosidyltrans
F;1-245/Product: coat protein 1A #status predicted <C1A>
F;246-491/Product: coat protein 1B #status predicted <C1B>
F;492-836/Product: coat protein 1C #status predicted <C1C>
F;837-980/Product: coat protein 2A #status predicted <C2A>
F;981-1076/Product: coat protein 2B #status predicted <C2B>
F;1077-1422/Product: coat protein 2C #status predicted <C2C>
F;1423-1484/Product: protein 3A #status predicted <C3A>
F;1485-1507/Product: protein 3B #status predicted <C3B>
F;1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F;1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
|||||
Db 834 FSQAKISLFYTEHEIMKFS 853

RESULT 4
GNNYMK
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core pro
NA polymerase (EC 2.7.7.48); protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: A94149; A25914; A94508
R;Cohen, J.I.; Rosenblum, B.; Tichhurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A;Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison wit
A;Reference number: A94149; MUID:87175701; PMID:3031686
A;Accession: A94149
A;Status: nucleic acid sequence not shown

A;Molecule type: genomic RNA
A;Residues: 1-2227 <COH>
A;Cross-references: EMBL:M16632; NID:G329594; PIDN:AAA45471.1; PID:G329595
A;Note: submitted to GenBank, August 1987
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleosidyltr
F;1-245/Product: coat protein 1A #status predicted <C1A>
F;246-491/Product: coat protein 1B #status predicted <C1B>
F;492-836/Product: coat protein 1C #status predicted <C1C>
F;837-980/Product: coat protein 2A #status predicted <C2A>
F;981-1076/Product: coat protein 2B #status predicted <C2B>
F;1077-1422/Product: coat protein 2C #status predicted <C2C>
F;1423-1484/Product: protein 3A #status predicted <C3A>
F;1485-1507/Product: protein 3B #status predicted <C3B>
F;1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F;1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
|||||
Db 834 FSQAKISLFYTEHEIMKFS 853

RESULT 5
GNNVHB
genome polyprotein - human hepatitis A virus (strain MBB)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
vpg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C;Accession: J50303
R;Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhardt,
Virus Res. 8, 153-171, 1987
A;Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isola
A;Reference number: J50303; MUID:86045071; PMID:2823500
A;Accession: J50303
A;Molecule type: genomic RNA
A;Residues: 1-2227 <PAU>
A;Cross-references: EMBL:M20273
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hyd
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-246/Product: coat protein 1B #status predicted <VP2>
F;247-491/Product: coat protein 1C #status predicted <VP3>
F;492-836/Product: coat protein 1D #status predicted <VP1>
F;837-980/Product: coat protein 2A #status predicted <P2A>
F;981-1108/Product: coat protein 2B #status predicted <P2B>
F;1109-1438/Product: coat protein 2C #status predicted <P2C>
F;1439-1496/Product: protein 3A #status predicted <P3A>
F;1497-1519/Product: genome-linked protein vpg #status predicted <VPG>
F;1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
|||||
Db 834 FSQAKISLFYTEHEIMKFS 853

RESULT 6
GNNYHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C;Accession: A03904
C;Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.
J. Virol. 54, 247-255, 1985
A;Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A;Reference number: A03904; MUID: 85185648; PMID: 2985793
A;Accession: A03904
A;Molecule type: genomic RNA
A;Residues: 1-852 <1IN>
A;Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593
C;Superfamily: hepatitis A virus genome polypeptide
C;Keywords: coat protein; core protein; polypeptide
F;1-245/Product: coat protein 1A #status predicted <1A>
F;246-491/Product: coat protein 1B #status predicted <1B>
F;492-836/Product: coat protein 1C #status predicted <1C>
F;837-852/Product: core protein 2A (fragment) #status predicted <2A>

Query Match 96.0%; Score 97; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKF 19
|||||
DB 834 FSOAKISLFYTEHEIMKF 852

RESULT 7
GNVSA
genome polypeptide - simian hepatitis A virus (strain AGM-27)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
C;Species: simian hepatitis A virus
C;Date: 31-Mar-1992 #sequence, revision 31-Mar-1992 #text_change 16-Jun-2000
C;Accession: A03470; S04885; S03965
R;Tsarev, S.A.
submitted to JIPiD, April 1991
A;Reference number: A03470
A;Accession: A03470
A;Molecule type: genomic RNA
A;Residues: 1-2230 <TSA>
A;Cross-references: GB:D00924; NID:G222597; PIDN:BAA00766.1; PID:G222598
R;Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A;Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an
A;Reference number: JQ1080; MUID:91311420; PMID:1649901
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhiko
submitted to the EMBL Data Library, May 1989
A;Reference number: S04885
A;Accession: S04885
A;Molecule type: genomic RNA
A;Residues: 1750-2164 <BAL2>
A;Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhiko
FEBS Lett. 247, 425-428, 1989
A;Title: Variations in genome fragments coding for RNA polymerase in human and simian he
A;Reference number: S03965; MUID:89232168; PMID:2541023
A;Accession: S03965
A;Molecule type: genomic RNA
A;Residues: 1960-2164 <BAL2>
A;Cross-references: EMBL:X15461
C;Superfamily: hepatitis A virus genome polypeptide
C;Keywords: coat protein; core protein; polypeptide
F;1-27/Product: coat protein 1A #status predicted <1A>
F;28-249/Product: coat protein 1B #status predicted <1B>
F;250-495/Product: coat protein 1C #status predicted <1C>
F;496-795/Product: coat protein 1D #status predicted <1D>
F;796-984/Product: core protein 2A #status predicted <2A>
F;985-1091/Product: core protein 2B #status predicted <2B>
F;1092-1426/Product: core protein 2C #status predicted <2C>
F;1427-1498/Product: protein 3A #status predicted <3A>
F;1499-1521/Product: protein 3B #status predicted <3B>
F;1522-1741/Product: protein 3C #status predicted <3C>
F;1742-2230/Product: protein 3D #status predicted <3D>

```

Query Match      87.1%; Score 88; DB 1; Length 2230;
Best Local Similarity 85.0%; Pred. No. 4.6e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FSQAKISLFYTEHEIMKFS 20
      |||||
Db      838 FSQAKISLFYTEHEIVKLS 857
      |||||

RESULT 8
F64300
formate dehydrogenase (EC 1.2.1.2) alpha chain - Methanococcus jannaschii
C/Species: Methanococcus jannaschii
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C/Accession: F64300
R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
R.; Reich, C.I.; Overbeek, R.; Kirkness, E.P.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
A.; Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A/Reference number: A64300; MUID:96337999; PMID:8688087
A/Accession: F64300
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-378 <BUL>
A/Cross-references: GB:U67459; GB:L77117; NID:G2826236; PIDN:AAB97987.1; PID:GI498763;
C/Genetics:
A/Map position: REV8474-7338
C/Keywords: oxidoreductase

Query Match      46.5%; Score 47; DB 2; Length 378;
Best Local Similarity 47.1%; Pred. No. 4.7;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy      4 AKISLFYTEHEIMKFS 20
      ::::
Db      174 SKVTIFNTEKEILKLN 190
      ::::

RESULT 9
H71869
hypothetical protein jhp0940 - Helicobacter pylori (strain J99)
C/Species: Helicobacter pylori
A/Variety: strain J99
C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
C/Accession: H71869
R/Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat
A/Reference number: A71800; MUID:99120557; PMID:9923682
A/Accession: H71869
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-325 <REN>
A/Cross-references: GB:AE001523; GB:AE001439; NID:G4155513; PIDN:AADC6516.1; PID:G41555
A/Experimental source: strain J99
C/Genetics:
A/Gene: jhp0940
A/Superfamily: Helicobacter pylori hypothetical protein jhp0940

```

```
Query Match      42.8%; Score 43; DB 2; Length 325;
Best Local Similarity 60.0%; Pred. NO. 19;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

RESULT 10

T22413
hypothetical protein F49C12.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22413
R:Gardner, A.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z19562
A:Accession: T22413
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-410 <WIL>
A:Cross-references: EMBL:Z68227; PIN:CAA92512.1; GSPDB:GN00022; CESP:F49C12.8
A:Experimental source: clone F49C12
C:Genetics:
A:Gene: CESP:F49C12.8
A:Map position: 4
A:Introns: 21/3; 302/3

Query Match 42.6%; Score 43; DB 2; Length 410;
Best Local Similarity 36.8%; Pred. NO. 24;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 FSQAISLYTTEHEIMKF 19
||::|::|
DB 156 FAMIRVGLFDDHHLINKF 174

RESULT 11
ABI1747
glycine cleavage system protein H homolog lin2519 [imported] - Listeria innocua (strain
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: ABI1747
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Arend, A.; Baquero, F.; Berche, P.; Bloecker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A.: Comparative genomics of Listeria species.
A:Reference number: AB1077; PMID:21537279; PMID:11679669
A:Accession: ABI1747
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <GLA>
A:Cross-references: GB:AL592022; PIN:CAC97746.1; PID:g16415041; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2519
C:Superfamily: glycine cleavage system protein H; lipoyl/biotin-binding homology

Query Match 41.6%; Score 42; DB 2; Length 125;
Best Local Similarity 72.7%; Pred. NO. 9.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LFYTEHEIMK 18
| | | | | | | |
DB 7 LLYTEHEWVK 17

RESULT 12
AI1377
glycine cleavage system protein H homolog lmo2425 [imported] - Listeria monocytogenes (S
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AI1377
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Anend, A.; Baquero, F.; Berche, P.; Bloecker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Poland, J.A.; Voss, H.; Wehlund,

A>Title: Comparative genomics of Listeria species.
A|Reference number: AB1077; MUID:21537279; PMID:11679669
A|Accession: A11377
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-125 <GLA>
A|Cross-references: GB|NC_003210; PIDN:CAD00503.1; PID:g16411913; GSPDB:GN00177
A|Experimental source: strain EGB-e
C|Genetics:
A|Gene: lmo2425
C|Superfamily: glycine cleavage system protein H; lipoyl/biotin-binding homology

Query Match 41.6%; Score 42; DB 2; Length 125;
Best Local Similarity 72.7%; Pred. NO. 9.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LEVTEHEIMK 18
| | | | | :
DB 7 LYTTEHEWVK 17
| | | | | :

RESULT 13
T18844
hypothetical protein C01H6.3 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C|Accession: T18844
R|Berk, M.
submitted to the EMBL Data Library, April 1996
A|Reference number: Z19030
A|Accession: T18844
A|Status: preliminary; translated from GB/EMBL/DDBJ
A|Molecule type: DNA
A|Residues: 1-187 <WIL>
A|Cross-references: EMBL:Z71258; PIDN:CAA95778.1; GSPDB:GN00019; CESP:C01H6.3
A|Experimental source: clone C01H6
C|Genetics:
A|Gene: CESP:C01H6.3
A|Map position: 1
A|Introns: 42/3; 82/1; 109/1
C|Superfamily: Caenorhabditis elegans hypothetical protein C01H6.3

Query Match 41.6%; Score 42; DB 2; Length 187;
Best Local Similarity 36.8%; Pred. NO. 15;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 FQAQKISLFYTEHEIMKF 19
| | | | | :
DB 161 FCNARISLQAQSHQCVR 179
| | | | | :

RESULT 14
S10842
gene ND5 intron 2 protein - Neurospora crassa mitochondrion
C|Species: Mitochondrion Neurospora crassa
C|Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 21-Jul-2000
C|Accession: S10842; S10081
R|Nelson, M.A.; Macino, G.
Mol. Gen. Genet. 206, 307-317, 1987
A|Title: Structure and expression of the overlapping ND4L and ND5 genes of Neurospora
A|Reference number: S07320; MUID:87228330; PMID:3035337
A|Accession: S10842
A|Molecule type: DNA
A|Residues: 1-545 <NEL>
A|Cross-references: EMBL:X05115
R|Almasan, A.; Mishra, N.C.
Genetics 120, 935-945, 1988
A|Title: Molecular characterization of the mitochondrial DNA of a new stopper mutant ER
A|Reference number: S10081; MUID:89137935; PMID:2976009
A|Accession: S10081
A|Molecule type: DNA
A|Residues: 457-545 <ALM>
A|Cross-references: EMBL:X14681; NID:g14021; PIDN:CAA32812.1; PID:g14022

Search completed: March 15, 2004, 14:06:58
Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:26:12 ; Search time 5.64706 Seconds
(without alignments)

184.415 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTERHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	1 POLG_HPAVH	P08617 hepatitis a
2	101	100.0	2227	1 POLG_HPAVL	P06441 hepatitis a
3	101	100.0	2227	1 POLG_HPAVM	P13901 hepatitis a
4	97	96.0	852	1 POLG_HPAVC	P06442 hepatitis a
5	93	92.1	2226	1 POLG_HPAV2	P26580 hepatitis a
6	93	92.1	2226	1 POLG_HPAV4	P26581 hepatitis a
7	93	92.1	2226	1 POLG_HPAV8	P26582 hepatitis a
8	88	87.1	2230	1 POLG_HPAVS	P14553 simian hepa
9	47	46.5	378	1 PDHA_METJA	G60314 methanococ
10	44	43.6	509	1 MATK_NYMOD	O98637 nympheae od
11	43	42.6	410	1 PSD6_CABEL	O20585 caenorhabdi
12	43	42.6	506	1 MATK_LATTI	O8mc8 lathyrus ti
13	43	42.6	722	1 YC06_KLEPN	O48452 klebsiella
14	42	41.6	125	1 GCSH_LISIN	Q92813 listeria in
15	42	41.6	125	1 GCSH_LISMO	Q9y412 listeria mo
16	41.5	41.1	288	1 NIF3_YEAST	P53081 saccharomyc
17	41	40.6	138	1 GCSH_PYPAB	O9v0g1 pyrococcus
18	41	40.6	304	1 META_THEMA	O9wz3 thermotoga
19	41	40.6	469	1 Y544_MYCPE	Q8evm0 mycoplasma
20	41	40.6	557	1 G6PI_ACICA	Q59088 acinetobact
21	41	40.6	1039	1 M2C1_MOUSE	Q91w89 mus musculu
22	41	40.6	1040	1 M2C1_RAT	P21139 rattus norv
23	41	40.6	2431	1 POLN_SPV	P08411 semliki for
24	40	39.6	134	1 GCSH_PYPFU	Q8u0u0 pyrococcus
25	40	39.6	418	1 B2AR_BOVIN	Q28044 bos taurus
26	40	39.6	5147	1 PCLO_HUMAN	O9y6v0 homo sapien
27	39.5	39.1	347	1 HTPX_PYRAE	O9zt30 pyrobaculum
28	39	38.6	152	1 Y765_METHH	O26859 methanobact
29	39	38.6	305	1 MK16_YEAST	P10962 saccharomyc
30	39	38.6	380	1 TRPD_YEAST	P07285 saccharomyc
31	39	38.6	534	1 UD15_HUMAN	P35504 homo sapien
32	39	38.6	634	1 YCX3_EUGGR	P31916 euglena gra
33	39	38.6	707	1 HLY2_ECOLI	P10089 escherichia

ALIGNMENTS

RESULT 1

POLG_HPAVH

ID POLG_HPAVH STANDARD; PRT; 2227 AA.

AC P08617; P06443; Q81082;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase

DE P3D (EC 2.7.7.48)]

OS Hepatitis A virus (strain HM-175).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12098;

RN [1]_TaxID=12098;

RP SEQUENCE FROM N.A.

RC STRAIN=Wild type;

RX MEDLINE=87061253; PubMed=3023706;

RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,

RA Baroudy B.M.;

RT "Complete nucleotide sequence of wild-type hepatitis A virus;

RT Comparison with different strains of hepatitis A virus and other

RT picornaviruses.";

RL J. Virol. 61:50-59(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Attenuated;

RX MEDLINE=87175701; PubMed=3031686;

RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,

RA Purcell R.H.;

RT "Complete nucleotide sequence of an attenuated hepatitis A virus;

RT Comparison with wild-type virus.";

RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).

RN [3]

RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.

RX MEDLINE=85166289; PubMed=2984684;

RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,

RA Purcell R.H., Feinstone S.M.;

RT "Sequence analysis of hepatitis A virus cDNA coding for capsid

RT proteins and RNA polymerase.";

RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).

CC - CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC {RNA}(N).

CC - SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC - PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC - MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED

CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.

CC - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC - CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT

CC SHOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; M14114; AAA45475.1; -
CC EMBL; M14707; AAA45465.1; -
CC EMBL; M14707; AAA45466.1; ALT_INIT.
CC EMBL; M16632; AAA45471.1; -
CC PIR; A03905; A03905.
CC PIR; A25981; GNNYHM.
CC PIR; A94149; GNNYMK.
CC PDB; 1HAV; 23-DEC-96.
CC INTERPRO; IPR004004; Calici_pol_hel.
CC INTERPRO; IPR009003; Cys Ser trypsin.
CC INTERPRO; IPR006095; RNA helicase.
CC INTERPRO; IPR007095; RNA_pol_DS_PS.
CC INTERPRO; IPR001205; RNA_pol_P3D.
CC INTERPRO; IPR007094; RNA_pol_P3D.
CC INTERPRO; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol_1.
CC Pfam; PF00910; RNA_helicase_1.
CC PRINTS; PR00918; CALICIVIRUSNS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT CHAIN 77 77
FT CHAIN 764 764
FT CHAIN 821 821
FT CHAIN 1052 1052
FT CHAIN 1062 1062
FT CHAIN 1118 1118
FT CHAIN 1151 1151
FT CHAIN 1163 1163
FT CHAIN 1277 1277
FT CHAIN 1500 1500
FT CHAIN 1805 1805
FT CHAIN 1930 1930
FT CHAIN 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853

```

RESULT 2

POLG_HPAVL

ID POLG_HPAVL STANDARD; PRT; 2227 AA.

AC P06441;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase

DE P3D (EC 2.7.7.48)].

OS Hepatitis A virus (strain 1A).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

```

OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Verwyweather J., van Nest G., Dina D.;
RT Primary structure and gene organization of human hepatitis A virus.;
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- FTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

EMBL; K02990; AAA45472.1; -

PIR; A03903; GNNYHR.

MEROPS; C03.005; -

InterPro; IPR004004; Calici_pol_hel.

InterPro; IPR009003; Cys Ser trypsin.

InterPro; IPR006095; RNA helicase.

InterPro; IPR007095; RNA_pol_DS_PS.

InterPro; IPR001205; RNA_pol_P3D.

InterPro; IPR007094; RNA_pol_P3D.

InterPro; IPR008975; Viral_cap_coat.

Pfam; PF00680; RNA_dep_RNA_pol_1.

Pfam; PF00910; RNA_helicase_1.

PRINTS; PR00918; CALICIVIRUSNS.

KW Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.

FT CHAIN 1 23

FT CHAIN 24 245

FT CHAIN 246 491

FT CHAIN 492 836

FT CHAIN 837 980

FT CHAIN 981 1076

FT CHAIN 1077 1422

FT CHAIN 1423 1484

FT CHAIN 1485 1507

FT CHAIN 1508 1678

FT CHAIN 1679 2227

FT CHAIN 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20

DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 3

POLG_HPAVL

ID POLG_HPAVL STANDARD; PRT; 2227 AA.

AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;

AC Q81090; Q81091; Q81092; Q81093;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase

DE P3D (EC 2.7.7.48)].

OS Hepatitis A virus (strain MEB).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Wiesel T., Kiehn R., Wimmer E.,
 RA Deinhardt F.;
 RT "The entire nucleotide sequence of the genome of human hepatitis A
 RT virus (isolate MB8).";
 RL Virus Res. 8:153-171(1987).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}[N].
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
 CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M20273; AAA45474.1; --
 DR MEROPS; C03.005; --
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006005; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_P5vir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
 FT CHAIN 837 980 CORE PROTEIN P2A.
 FT CHAIN 981 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
 FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
 Query Match 100.0%; Score 101; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSQAKISLFYTEHEIMKFS 20
 DB 834 FSQAKISLFYTEHEIMKFS 853
 RESULT 4
 POLG_HPAVC
 ID POLG_HPAVC STANDARD; PRT; 852 AA.
 AC P06442; Q83741; Q83742;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein

DE P2A] (Fragment).
 OS Hepatitis A virus (strain CR326).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85185648; PubMed=2985793;
 RA Linenmeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
 RA Young A., Mitra S.W.;
 RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
 RL J. Virol. 54:247-255(1985).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M10033; AAA45470.1; --
 DR PIR; A03904; GNNYHA.
 DR InterPro; IPR008975; Viral_cap_coat.
 KW Polyprotein; Coat protein; Core protein.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
 FT CHAIN 837 852 CORE PROTEIN P2A.
 FT NON_TER 852 852
 SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;
 Query Match 96.0%; Score 97; DB 1; Length 852;
 Best Local Similarity 100.0%; Pred. No. 4e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSQAKISLFYTEHEIMKFS 19
 DB 834 FSQAKISLFYTEHEIMKFS 852
 RESULT 5
 POLG_HPAV2
 ID POLG_HPAV2 STANDARD; PRT; 2226 AA.
 AC P26580;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain 24a).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12094;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705595;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
 RA Cromeans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}[N].
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PFM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M59810; AAA45469.1; --
CC MEROPS: C03.005; --
CC InterPro: IPR004004; Calici_pol_hel.
CC InterPro: IPR009003; Cys Ser trypsin.
CC InterPro: IPR006005; RNA_helicase.
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR007094; RNA_pol_PSVir.
CC InterPro: IPR008975; Viral_cap_coat.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC PRINTS: PR00918; CALICIVIRUS.
CC Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
KW CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. NO. 5e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOANISLFYTEHEIMKFS 853

RESULT 6
POLG HPVA4 STANDARD; PRT; 2226 AA.
ID POLG HPVA4
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12095;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91162758; PubMed=1705995;
RA Lemmon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RL J. Virol. 65:2056-2065 (1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

(RNA)(N).
-1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
-1- PFM: Specific enzymatic cleavages in vivo yield mature proteins.
-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
EMBL: M59809; AAA45469.1; --
MEROPS: C03.005; --
InterPro: IPR004004; Calici_pol_hel.
InterPro: IPR009003; Cys Ser trypsin.
InterPro: IPR006005; RNA_helicase.
InterPro: IPR007095; RNA_pol_DS_PS.
InterPro: IPR001205; RNA_pol_P3D.
InterPro: IPR007094; RNA_pol_PSVir.
InterPro: IPR008975; Viral_cap_coat.
Pfam: PF00680; RNA_dep_RNA_pol; 1.
Pfam: PF00910; RNA_helicase; 1.
PRINTS: PR00918; CALICIVIRUS.
Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
KW CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9EF75 CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. NO. 5e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOANISLFYTEHEIMKFS 853

RESULT 7
POLG HPVA8 STANDARD; PRT; 2226 AA.
ID POLG HPVA8
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12096;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91162758; PubMed=1705995;
RA Lemmon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RT

```

RL J. Virol. 65:2056-2065(1991).
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL; M59808; AAA45467.1; -.
CC PDB; 1QA7; 15-MAY-00.
CC MEROPS; C03.005; -.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. No. 5e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 834 FSQAKISLFYTEHEIMKFS 853

RESULT 8
POLG.HPAVS STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12102;
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;

```

```

RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1693(1991).
RN [2]
RX SEQUENCE OF 1750-2164 FROM N.A.
RA MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kurov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL; D00924; BAA00766.1; -.
CC EMBL; X15461; CAA33490.1; -.
CC PIR; A30470; GNNYSA.
CC MEROPS; C03.005; -.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 87.1%; Score 88; DB 1; Length 2230;
Best Local Similarity 85.0%; Pred. No. 3.3e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 838 FSQAKISLFYTEHEIMKFS 857

RESULT 9
FDHA.METJA STANDARD; PRT; 378 AA.
AC Q60314;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase alpha chain (EC 1.2.1.2).

```


GN MJ0006.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.E., Fuhmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurd M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL Science 273:1058-1073(1996).
CC -|- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -|- COFACTOR: Molybdenum (molybdopterin), zinc, FAD. May bind a 4Fe-4S
CC cluster (By similarity).
CC -|- SUBUNIT: Dimer of alpha and beta chains (By similarity).
CC -|- SIMILARITY: Belongs to the prokaryotic molybdopterin-containing
CC oxidoreductase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U67459; AAB97987.1; -;
CC F1R; F64300; F64300.
CC HSP; P07658; 1AA6.
CC TIGR: MJ0006; -;
CC InterPro: IPR006963; Molybdop Fe4S4.
CC InterPro: IPR006656; Molybdopterin.
CC InterPro: IPR006655; Prok_Moxred.
CC Pfam: PF04879; Molybdop Fe4S4; 1.
CC Pfam: PF00384; molybdopterin; 2.
CC PROSITE: PS00551; MOLYBDOPTERIN_PROK_1; 1.
CC PROSITE: PS00490; MOLYBDOPTERIN_PROK_2; 1.
CC PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; FALSE NEG.
CC Hypothetical protein; Oxidoreductase; zinc; Flavoprotein; Molybdenum;
KW FAD; NAD; iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 8 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 11 11 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 15 15 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 43 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 378 AA; 42050 MW; E0C13060FF2609AA CRC64;

Query Match 46.5%; Score 47; DB 1; Length 378;
Best Local Similarity 47.1%; Pred. No. 2.9;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 AKISLFYTEEHEIMKFS 20
Db 174 SKVTIFNTEKEIKLN 190

RESULT 10
MATK_NYMOD
ID MATK_NYMOD STANDARD; PRT; 509 AA.
AC Q98637;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Maturase K (intron maturase).
GN MATK.

OS Nymphaea odorata (White water lily).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; stem Magnoliophyta; Nymphaeaceae;
OC Nymphaea.
OX NCBI_TaxID=4419;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee D.H., Schneider E.L., Padgett D.J., Soltis P.S., Soltis D.E.,
RA Zanis M.;
RT "Phylogeny, classification and floral evolution of water lilies
RT (Nymphaeaceae; Nymphaeales): a synthesis of non-molecular, rbcL, matK
RT and 18S rDNA data";
RL Syst. Bot. 24:28-46(1999).
CC -|- FUNCTION: Probably assists in splicing chloroplast group II
CC introns (By similarity).
CC -|- SIMILARITY: Belongs to the intron maturase family 2. MatK
CC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AF092988; AAD05556.1; -;
CC InterPro: IPR000442; Intron_maturase2.
CC InterPro: IPR002866; MatK_N.
CC Pfam: PF01348; Intron_maturase2; 1.
CC Pfam: PF01824; MatK_N; 1.
KW mRNA processing; Chloroplast.
SQ SEQUENCE 509 AA; 60343 MW; C78FCA7301BDF373 CRC64;

Query Match 43.6%; Score 44; DB 1; Length 509;
Best Local Similarity 72.7%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 9 FYTEEHEIMKF 19
Db 472 FLTEEHEIVSF 482

RESULT 11
PSD6_CAEEL
ID PSD6_CAEEL STANDARD; PRT; 410 AA.
AC Q20585; 1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable 26S proteasome non-ATPase regulatory subunit 6.
GN RPN-7 OR P49C12.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: Acts as a regulatory subunit of the 26S proteasome which
CC is involved in the ATP-dependent degradation of ubiquitinated
CC proteins (By similarity).
CC -|- SIMILARITY: Contains 1 PCI domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68227; CAA92512.1; -.
DR PIR; T22413; T22413
DR WormPep; F49C12.8; CE03368.
DR InterPro; IPR000717; PCI.
DR Pfam; PF01399; PCI; 1.
DR SMART; SMO0088; PINT; 1.
KW Proteasome.
FT DOMAIN 205 379 PCI.
SQ SEQUENCE 410 AA; 47583 MW; F37390A3250109EE CRC64;

Query Match 42.6%; Score 43; DB 1; Length 410;
Best Local Similarity 36.8%; Pred. No. 14;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMK 19
Db 156 FAMIRVGLFLLDHLINKF 174

RESULT 12
MATK LATTI
ID MATK LATTI STANDARD; PRT; 506 AA.
AC Q8MCR8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Maturase K (Intron maturase).
GN MATK.
OS Lathyrus tingitanus (Tangier pea).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Lathyrus.
OX NCBI_TaxID=3862;
RN [1]
RP SEQUENCE FROM N.A.
RA Steele K.P., Wojciechowski M.F.;
RT "Phylogenetic analyses of tribes Trifolieae and Viciae based on
sequences of the plastid gene matK (Papilionoideae: Leguminosae).";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: Probably assists in splicing chloroplast group II
introns (By similarity).
CC -!- SIMILARITY: Belongs to the intron maturase family 2. MatK
subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF522087; AAM82079.1; -.
DR InterPro; IPR000442; Intron_maturase2.
DR InterPro; IPR002866; MatK_N.
DR Pfam; PF01348; Intron_maturase2; 1.
DR Pfam; PF01824; MatK_N; 1.
KW mRNA processing; Chloroplast.
SQ SEQUENCE 506 AA; 60897 MW; 6187B6AE73C9D564 CRC64;

Query Match 42.6%; Score 43; DB 1; Length 506;
Best Local Similarity 44.4%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 SQAKISLFYTEHEIMK 19
Db 457 SEELQEFTEQEILSF 474

RESULT 13

```

```

YC06 KLEPN
ID YC06 KLEPN STANDARD; PRT; 722 AA.
AC Q484E2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative tyrosine-protein kinase in cps region (EC 2.7.1.112) (ORF6).
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chedi;
RX MEDLINE=95204345; PubMed=7896702;
RA Arakawa Y., Wacharotayankun R., Nagatsuka T., Ito H., Kato N.,
Ohta M.;
RT "Genomic organization of the Klebsiella pneumoniae cps region
responsible for serotype K2 capsular polysaccharide synthesis in the
virulent strain Chedi";
RL J. Bacteriol. 177:1788-1796(1995).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -!- PATHWAY: Exopolysaccharide biosynthesis.
CC -!- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ETX/WZC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D21242; BAA04777.1; -.
DR InterPro; IPR003856; LPS_WZZ_MPA.
DR Pfam; PF02706; wzz; 1.
KW Hypothetical protein; Transferase; Tyrosine-protein kinase;
KW Exopolysaccharide synthesis; Transmembrane; Inner membrane;
KW ATP-binding.
FT TRANSMEM 31 53 POTENTIAL.
FT TRANSMEM 427 449 POTENTIAL.
SQ SEQUENCE 722 AA; 80400 MW; 3CAD6910AE81C3D7 CRC64;

Query Match 42.6%; Score 43; DB 1; Length 722;
Best Local Similarity 44.4%; Pred. No. 26;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMK 18
Db 326 FRESEISQLYTKHEPTYK 343

RESULT 14
GCSH LISIN
ID GCSH LISIN STANDARD; PRT; 125 AA.
AC Q928L3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glycine cleavage system H protein.
GN GCVH OR LIN2519.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
Baquero A., Berche P., Bloeker H., Brandt P., Chakraborty T.,
Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

```

RA	Donann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA	Entian K.-D., Fshi H., Garcia-del Portillo F., Garrido P.,
RA	Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA	Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA	Madueno E., Maicounam A., Mata Vicente J., Ng E., Nedjari H.,
RA	Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA	Rommel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA	Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RA	"Comparative genomics of <i>Listeria species</i> ."
RA	Science 294:849-852(2001).
CC	-!- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC	glycine. The H protein shuttles the methylamine group of glycine
CC	from the P protein to the T protein (By similarity).
CC	-!- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
CC	(By similarity).
CC	-!- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC	P, T, L and H (By similarity).
CC	-!- SIMILARITY: Belongs to the gcvH family.
CC	-!- SIMILARITY: Contains 1 lipoyl-binding domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AL596172; CAC97746.1; -
DR	PIR; AB1747; AB1747.
DR	Listlist; LIN02519; -
DR	HAMAP; MF_00272; -; 1.
DR	InterPro; IPR002930; GCV_H.
DR	InterPro; IPR003016; Lipoyl_BS.
DR	Pfam; PF01597; GCV_H; 1.
DR	TIGRFAMs; TIGR00527; GCVH; 1.
DR	PROSITE; PS00189; LIPOYL; 1.
KW	Lipoyl; Complete proteome.
FT	BINDING 63
FT	53 LIPOYL (BY SIMILARITY).
SS	SEQUENCE 125 AA; 13732 MW; 94F00032FECFA152 CRC64;
SS	Query Match 41.6%; Score 42; DB 1; Length 125;
SS	Best Local Similarity 72.7%; Pred. No. 6.1;
SS	Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY	8 LFYTEEHEIMK 18
DB	1 :
DB	7 LLYTEEHEWK 17
RESULT 15	
GCSD	GCSD
ID	GCSD_LISMO STANDARD; PRT; 125 AA.
AC	Q8Y4L2;
DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DE	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Glycine cleavage system H protein.
GN	GCVH OR LMC2425.
OS	Listeria monocytogenes.
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX	NCBI_TaxID=1639;
RN	[1]_TaxID=1639;
RP	SEQUENCE FROM N.A.
RC	STRAIN=EGD-e / Serovar 1/2a;
RC	MEDLINE=21537279; PubMed=11679669;
RA	Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA	Baquero F., Berche P., Blecker H., Brandt P., Chakraborty T.,
RA	Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA	Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA	Entian K.-D., Fshi H., Garcia-del Portillo F., Garrido P.,
RA	Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA	Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,

```

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordziel G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tlerriz A.,
RA Vazquez-Bolland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC glycine. The H protein shuttles the methylamine group of glycine
CC from the P protein to the T protein (By similarity).
CC -!- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
CC (By similarity).
CC -!- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC P, T, L and H (By similarity).
CC -!- SIMILARITY: Belongs to the gcvH family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL591983; CAD00503.1; -
DR PIR; A11377; A11377.
DR Listlist; LMO02425; -.
DR HAMAP; MF_00272; -.
DR InterPro; IPR002930; GCV_H.
DR InterPro; IPR003016; Lipoyl_BS.
DR Pfam; PF01597; GCV_H; 1
DR TIGRFAMs; TIGR00527; gcvH; 1.
DR PROSITE; PSC0189; LIPOYL; 1.
DR Lipoyl; Complete proteome.
DR BINDING 63 63
DR FT BINDING 63 63 LIPOYL (BY SIMILARITY).
DR SEQUENCE 125 AA; 13801 MW; 1B4EE1E37C8980B5 CRC64;
SQ
Query Watch 41.6%; Score 42; DB 1; Length 125;
Best Local Similarity 72.7%; Pred. No. 6.1;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 8 LEVTEEHEIMK 18
DB 7 LLYTEHEWVK 17

```

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:49:12 ; Search time 29.1765 Seconds
(without alignments)
216.283 Million cell updates/sec

Title: US-09-171-432A-43
Perfect score: 101
Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP TREMBL 25:*

2: sp_archaea:*

3: sp_bacteria:*

4: sp_fungi:*

5: sp_human:*

6: sp_invertebrate:*

7: sp_mammal:*

8: sp_mmc:*

9: sp_organelle:*

10: sp_phase:*

11: sp_plant:*

12: sp_rodent:*

13: sp_virus:*

14: sp_vertebrate:*

15: sp_unclassified:*

16: sp_rvirus:*

17: sp_bacteriap:*

18: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	112	12 Q86534	Q86534 hepatitis a
2	101	100.0	132	12 Q8V4L9	Q8V4L9 hepatitis a
3	101	100.0	132	12 Q8V4L6	Q8V4L6 hepatitis a
4	101	100.0	132	12 Q8V4N9	Q8V4N9 hepatitis a
5	101	100.0	132	12 Q8V4T9	Q8V4T9 hepatitis a
6	101	100.0	132	12 Q8V4N1	Q8V4N1 hepatitis a
7	101	100.0	132	12 Q8V4J5	Q8V4J5 hepatitis a
8	101	100.0	132	12 Q8V4M5	Q8V4M5 hepatitis a
9	101	100.0	132	12 Q8V4K9	Q8V4K9 hepatitis a
10	101	100.0	132	12 Q8V4N3	Q8V4N3 hepatitis a
11	101	100.0	132	12 Q8V4M7	Q8V4M7 hepatitis a
12	101	100.0	132	12 Q8V4N2	Q8V4N2 hepatitis a
13	101	100.0	132	12 Q8V4K7	Q8V4K7 hepatitis a
14	101	100.0	132	12 Q8V4K6	Q8V4K6 hepatitis a
15	101	100.0	132	12 Q8V4M4	Q8V4M4 hepatitis a
16	101	100.0	132	12 Q8V4N6	Q8V4N6 hepatitis a

17	101	100.0	132	12 Q8V4J7	Q8V4J7 hepatitis a
18	101	100.0	132	12 Q8V4L5	Q8V4L5 hepatitis a
19	101	100.0	132	12 Q8V4K2	Q8V4K2 hepatitis a
20	101	100.0	132	12 Q8V4K3	Q8V4K3 hepatitis a
21	101	100.0	132	12 Q8V4N7	Q8V4N7 hepatitis a
22	101	100.0	132	12 Q8V4K1	Q8V4K1 hepatitis a
23	101	100.0	132	12 Q8V4J8	Q8V4J8 hepatitis a
24	101	100.0	132	12 Q8V4L0	Q8V4L0 hepatitis a
25	101	100.0	132	12 Q8V4K4	Q8V4K4 hepatitis a
26	101	100.0	132	12 Q8V4M2	Q8V4M2 hepatitis a
27	101	100.0	132	12 Q8V4M3	Q8V4M3 hepatitis a
28	101	100.0	132	12 Q8V4L8	Q8V4L8 hepatitis a
29	101	100.0	132	12 Q8V4N5	Q8V4N5 hepatitis a
30	101	100.0	132	12 Q8V4N4	Q8V4N4 hepatitis a
31	101	100.0	132	12 Q8V4K8	Q8V4K8 hepatitis a
32	101	100.0	132	12 Q8V4M6	Q8V4M6 hepatitis a
33	101	100.0	132	12 Q8V4M8	Q8V4M8 hepatitis a
34	101	100.0	132	12 Q8V4L2	Q8V4L2 hepatitis a
35	101	100.0	132	12 Q8V4K0	Q8V4K0 hepatitis a
36	101	100.0	132	12 Q8V4L7	Q8V4L7 hepatitis a
37	101	100.0	132	12 Q8V4M1	Q8V4M1 hepatitis a
38	101	100.0	132	12 Q8V4M0	Q8V4M0 hepatitis a
39	101	100.0	132	12 Q8V4K5	Q8V4K5 hepatitis a
40	101	100.0	155	12 Q8B8M5	Q8B8M5 hepatitis a
41	101	100.0	155	12 Q8B8M4	Q8B8M4 hepatitis a
42	101	100.0	155	12 Q8B8M3	Q8B8M3 hepatitis a
43	101	100.0	155	12 Q8B8M2	Q8B8M2 hepatitis a
44	101	100.0	155	12 Q8B8M1	Q8B8M1 hepatitis a
45	101	100.0	155	12 Q8B8M0	Q8B8M0 hepatitis a

ALIGNMENTS

RESULT 1

Q86534 PRELIMINARY; PRT; 112 AA.
AC Q86534;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VPI/2A (Fragment).
GN VPI/2A.
OS Hepatitis A virus.
OC Viruses; serona positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSH/S;
RX MEDLINE=92348853; PubMed=1668326;
RA Pineschi N., Cavalieri F., Garelick H., Prugnola A., Pellegrini V.,
RA Zuckerman A.J.;
RT "Characterization of a hepatitis A virus strain suitable for vaccine production."
RL J. Hepatol. 13:S146-S151(1991).
DR EMBL; S44105; AAB2739.2; -;
FT NON TER 112 112
SQ SEQUENCE 112 AA; 13145 MW; 3C0CBBB4570D9A2C CRC64;

Query Match 100.0%; Score 101; DB 12; Length 112;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20

Db 75 FSQAKISLFYTEHEIMKFS 94

RESULT 2

Q8V4L9 PRELIMINARY; PRT; 132 AA.
ID Q8V4L9
AC Q8V4L9;

DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA-11/07/00;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386864; AAL68525.1; -.
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 92 FSQAKISLFYTEHEIMKFS 111

RESULT 3
Q8V4L6 PRELIMINARY; PRT; 132 AA.
AC Q8V4L6
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA-01/09/00;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386867; AAL68528.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 92 FSQAKISLFYTEHEIMKFS 111

RESULT 4
Q8V4M9 PRELIMINARY; PRT; 132 AA.
ID Q8V4M9

AC Q8V4M9;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA-19/11/97;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386854; AAL68515.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
FT NON_TER 1
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 92 FSQAKISLFYTEHEIMKFS 111

RESULT 5
Q8V4J9 PRELIMINARY; PRT; 132 AA.
AC Q8V4J9
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-06/04/99;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386884; AAL68545.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 92 FSQAKISLFYTEHEIMKFS 111

RESULT 6
Q8V4N1 PRELIMINARY; PRT; 132 AA.
ID Q8V4N1

AC Q8V4N1;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA-17/06/97;
 RX MEDLINE=21571641; PubMed=11714971;
 RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
 RT "Genetic analysis of hepatitis A virus strains recovered from the
 environment and from patients with acute hepatitisis.";
 RL J. Gen. Virol. 82:2955-2963(2001).
 DR EMBL; AF386852; AAL68513.1; -.
 DR PIR; PQ0427; PQ0428.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 15354 MW; D05770901936C69D CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 132;
 Best Local Similarity 100.0%; Pred. No. 6.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSQAKISLFYTEHEIMKFS 20
 |||||
 Db 92 FSQAKISLFYTEHEIMKFS 111
 |||||
 RESULT 7
 Q8V4J5
 ID Q8V4J5 PRELIMINARY; PRT; 132 AA.
 AC Q8V4J5;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA-21/11/00;
 RX MEDLINE=21571641; PubMed=11714971;
 RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
 RT "Genetic analysis of hepatitis A virus strains recovered from the
 environment and from patients with acute hepatitisis.";
 RL J. Gen. Virol. 82:2955-2963(2001).
 DR EMBL; AF386858; AAL68549.1; -.
 DR PIR; PQ0427; PQ0427.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 132;
 Best Local Similarity 100.0%; Pred. No. 6.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSQAKISLFYTEHEIMKFS 20
 |||||
 Db 92 FSQAKISLFYTEHEIMKFS 111
 |||||
 RESULT 8
 Q8V4M5
 ID Q8V4M5 PRELIMINARY; PRT; 132 AA.
 AC Q8V4M5;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA-11/02/98;
 RX MEDLINE=21571641; PubMed=11714971;
 RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
 RT "Genetic analysis of hepatitis A virus strains recovered from the
 environment and from patients with acute hepatitisis.";
 RL J. Gen. Virol. 82:2955-2963(2001).
 DR EMBL; AF386858; AAL68513.1; -.
 DR PIR; PQ0427; PQ0427.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 132;
 Best Local Similarity 100.0%; Pred. No. 6.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSQAKISLFYTEHEIMKFS 20
 |||||
 Db 92 FSQAKISLFYTEHEIMKFS 111
 |||||

AC Q8V4M5;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA-11/02/98;
 RX MEDLINE=21571641; PubMed=11714971;
 RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
 RT "Genetic analysis of hepatitis A virus strains recovered from the
 environment and from patients with acute hepatitisis.";
 RL J. Gen. Virol. 82:2955-2963(2001).
 DR EMBL; AF386858; AAL68513.1; -.
 DR PIR; PQ0427; PQ0427.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0430; PQ0430.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 15284 MW; D8FC83D4BC38287D CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 132;
 Best Local Similarity 100.0%; Pred. No. 6.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSQAKISLFYTEHEIMKFS 20
 |||||
 Db 92 FSQAKISLFYTEHEIMKFS 111
 |||||
 RESULT 9
 Q8V4K9
 ID Q8V4K9 PRELIMINARY; PRT; 132 AA.
 AC Q8V4K9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA-05/03/90;
 RX MEDLINE=21571641; PubMed=11714971;
 RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
 RT "Genetic analysis of hepatitis A virus strains recovered from the
 environment and from patients with acute hepatitisis.";
 RL J. Gen. Virol. 82:2955-2963(2001).
 DR EMBL; AF386874; AAL68535.1; -.
 DR PIR; PQ0427; PQ0427.
 DR PIR; PQ0428; PQ0428.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 132;
 Best Local Similarity 100.0%; Pred. No. 6.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSQAKISLFYTEHEIMKFS 20
 |||||
 Db 92 FSQAKISLFYTEHEIMKFS 111
 |||||
 RESULT 10
 Q8V4N3

```
ID Q8V4N3 PRELIMINARY; PRT; 132 AA.
AC Q8V4N3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11-18/12/95;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RL environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386850; AAL68511.1; -.
DR PIR; P0427; P0427.
DR PIR; P0428; P0428.
DR PIR; P0431; P0431.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15199 MW; B0C6D1A19DCF3AF2 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB |||||
92 FSOAKISLFYTEHEIMKFS 111

RESULT 11
Q8V4M7 PRELIMINARY; PRT; 132 AA.
AC Q8V4M7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA-16/01/98;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RL environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386856; AAL68517.1; -.
DR PIR; P0427; P0427.
DR PIR; P0428; P0428.
DR PIR; P0430; P0430.
DR PIR; P0431; P0431.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15284 MW; D8FC83D4BC38287D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB |||||
92 FSOAKISLFYTEHEIMKFS 111

RESULT 12
Q8V4N2 PRELIMINARY; PRT; 132 AA.
AC Q8V4N2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA-12/02/96;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RL environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386851; AAL68512.1; -.
DR PIR; P0427; P0427.
DR PIR; P0428; P0428.
DR PIR; P0430; P0430.
DR PIR; P0431; P0431.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15268 MW; D8FC83D50C31B87D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB |||||
92 FSOAKISLFYTEHEIMKFS 111

RESULT 13
Q8V4K7 PRELIMINARY; PRT; 132 AA.
AC Q8V4K7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-17/01/91;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RL environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386876; AAL68537.1; -.
DR PIR; P0427; P0427.
DR PIR; P0428; P0428.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB |||||
92 FSOAKISLFYTEHEIMKFS 111
```

Search completed: March 15, 2004, 14:05:21
Job time : 30.1765 secs

RESULT 14
QSV4K6 PRELIMINARY; PRT; 132 AA.
AC QSV4K6; 20, Created
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-27/02/91;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jordi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386877; AAL68538.1; -.
DR PIR; PQ0427;
DR PIR; PQ0428;
DR PIR; PQ0430;
DR PIR; PQ0431;
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15266 MW; D8FC83D4B97D3C29 CRC64;
Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20
DB 92 FSQAKISLFYTEHEIMKFS 111

RESULT 15
QSV4M4 PRELIMINARY; PRT; 132 AA.
AC QSV4M4; 20, Created
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA-22/03/99;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jordi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386859; AAL68520.1; -.
DR PIR; PQ0427;
DR PIR; PQ0428;
DR PIR; PQ0431;
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;
Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKFS 20

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:24:17 ; Search time 44.2353 Seconds
(without alignments)
127.748 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KVNPHGMLEEEAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_28Jan04: *
1: Genesep1980s: *
2: Genesep1990s: *
3: Genesep2000s: *
4: Genesep2001s: *
5: Genesep2002s: *
6: Genesep2003as: *
7: Genesep2003bs: *
8: Genesep2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	20	2	AAW42927
2	104	100.0	21	4	AAB69444 Synthetic
3	104	100.0	2227	1	AAP60066 Sequence
4	104	100.0	2227	2	AAR05697 Attenuate
5	104	100.0	2227	2	AAW34074 Hepatitis
6	104	100.0	2227	3	AAB18609 Amino aci
7	104	100.0	2227	3	AAB18607 Amino aci
8	104	100.0	2227	3	AAB18608 Amino aci
9	104	100.0	2227	5	AAE19899 Hepatitis
10	104	100.0	2227	5	ABG31729 Attenuate
11	104	100.0	2227	5	ABG31727 Wild-type
12	104	100.0	2227	5	ABG31728 Hepatitis
13	104	100.0	2227	6	ABU08640 Attenuat
14	104	100.0	2227	6	ABU08641 Attenuat
15	104	100.0	2227	6	ABU08639 Wild type
16	104	100.0	2227	7	ABW00350 Hepatitis
17	53	51.0	20	2	AAW42928 Immunogen
18	53	51.0	21	4	AAB69445 Synthetic
19	47	45.2	387	5	ABBS3661 Lactococ
20	46	44.2	259	6	ABU49511 Protein e
21	46	44.2	444	6	ABM73085 Staphyloc
22	45	43.3	195	4	AAB32447 Novel hum
23	45	43.3	206	3	AAG32341 Arabidops
24	45	43.3	242	3	AAG32340 Arabidops
25	45	43.3	242	3	AAG10097 Arabidops

26	45	43.3	383	3	AAG10096 Arabidops
27	45	43.3	383	3	AAG32339 Arabidops
28	45	43.3	399	3	AAG10095 Arabidops
29	44	42.3	102	4	ABG11781 Novel hum
30	44	42.3	295	6	ABU43671 Protein e
31	44	42.3	432	2	AAV49167 Human SCS
32	44	42.3	432	3	AAV76886 Human SCS
33	44	42.3	432	4	AAU32446 Novel hum
34	44	42.3	500	6	ABR53060 Protein s
35	43.5	41.8	267	3	AAG37736 Arabidops
36	43.5	41.8	267	3	AAG37721 Arabidops
37	43.5	41.8	269	3	AAG24210 Arabidops
38	43.5	41.8	426	3	AAG37735 Arabidops
39	43.5	41.8	426	3	AAG37720 Arabidops
40	43.5	41.8	426	5	ABB93930 Herbicida
41	43.5	41.8	428	3	AAG24209 Arabidops
42	43.5	41.8	438	3	AAG37734 Arabidops
43	43.5	41.8	440	3	AAG24208 Arabidops
44	43.5	41.8	442	3	AAG37719 Arabidops
45	43	41.3	98	2	AAW29645 Human sec

ALIGNMENTS

RESULT 1

AAW42927
ID AAW42927 standard; peptide; 20 AA.

XX AC AAW42927;

DT 28-APR-1998 (first entry)

DE Immunogenic Hepatitis A virus peptide YK-1327.

XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.

XX Synthetic.

OS Hepatitis A virus.

XX WO9740147-A1.

PD 30-OCT-1997.

XX 18-APR-1997; 97WO-US006891.

XX 19-APR-1996; 96US-0015644P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal.

XX Claim 18; Page 112; 140pp; English.

PS Peptides AAW42922-30 are immunogenic peptides corresponding to immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are substantially similar to a portion of the amino acid sequence of the P2A protein of HAV corresponding to amino acids 922-941, and has a reactivity of 31.3% with acute sera. Compositions containing the peptides can be used to induce an immune response to HAV in a mammal. The peptides can also be used to detect the presence of antibodies against HAV in mammalian serum. The peptides can also be used to make an antibody against HAV by administering the peptide to a mammal

XX Sequence 20 AA;

Query Match 100.0%; Score 104; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIEAANSKD 20
 |||||
 DB 1 KVNPHGMLDLEIEAANSKD 20

RESULT 2
 AAB69444
 ID AAB69444 standard; peptide; 21 AA.

AC AAB69444;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 44.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 XX antigen; major structural capsid polypeptide; HAV antibody detection.
 OS
 OS Hepatitis A virus.
 OS Synthetic.

PN WO200105824-A2.
 XX
 PD 25-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-USO19267.
 XX
 PR 15-JUL-1999; 99US-0144412P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE;
 XX
 DR WPI; 2001-112681/12.
 XX

Synthetic peptides used as antigen sources for enzyme immunoassays
 detecting anti-hepatitis A virus and as vaccines.
 XX
 PF Claim 13; Page 96; 130pp; English.
 XX

The present sequence is one of a number of synthetic peptides which are
 immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 comprise antigenic epitopes of the major structural capsid polypeptides
 or non-structural polypeptides of HAV with one or more glutamine
 molecules at the carboxy end of the peptide. The peptides are used to
 detect the presence of antibodies against HAV in mammalian serum, to
 detect the presence of HAV in a human or animal through the binding of
 the peptide to an antibody, to detect acute phase infection by detecting
 IGM antibodies in mammalian serum and detecting convalescence in a
 mammal. The peptides are used to detect or quantify HAV antibodies in
 samples in clinical or research-based assays using immunoblotting,
 fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 tracking of radioactive or bioluminescent markers, chromatography or
 electrophoresis. The peptides are used to induce an immune response to
 HAV when administered to a human or animal. Glutamine at the carboxy end
 of the peptides enhances the IGM antibody reactivity
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 104; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.3e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIEAANSKD 20
 |||||
 DB 1 KVNPHGMLDLEIEAANSKD 20

RESULT 3

AAP60066

ID AAP60066 standard; protein; 2227 AA.

AC AAP60066;
 XX
 DT 25-MAR-2003 (revised)
 DT 26-JUN-1991 (first entry)
 XX
 DE Sequence of viral I434 polypeptide encoded by the complete nucleotide
 DE sequence of the HAV genome.
 XX
 KW Diagnosis; vaccine; passive immunotherapy.
 XX
 OS Hepatitis A virus.

XX
 FH Key Location/Qualifiers
 FT Region 1..245
 FT /label= P1.1A
 FT Region 246..491
 FT /label= 1B
 FT Region 492..836
 FT /label= 1C
 FT Region 837..980
 FT /label= P2.2A
 FT Region 981..1076
 FT /label= 2B
 FT Region 1077..1422
 FT /label= 2C
 FT Region 1423..1484
 FT /label= P3.3A
 FT Region 1485..1507
 FT /label= 3B
 FT Region 1508..1678
 FT /label= 3C
 FT Region 1679..2227
 FT /label= 3D
 XX
 PN EP199480-A.

XX

PD 29-OCT-1986.

XX

PF 03-APR-1986; 86EP-00302465.

XX

PR 03-APR-1985; 85US-00719329.

XX

PA (CHIR) CHIRON CORP.

XX

PI Dina D, Potter SJ, Vannest GA, Caput D;

XX

PI WPI; 1986-286213/44.

XX

DR N-PSDB; AAN60080.

XX

PT Hepatitis A virus nucleotide sequence and polypeptide - and use in prodn.

XX

PS Claim 5; Fig 1; 18pp; English.

XX

AAAN60080 and oligonucleotide fragments are useful in detection of

CC hepatitis A virus; transformed hosts may be used for expression of

CC polypeptides and fragments useful in vaccines without risk of infection

CC by the virus or in prodn. of particles which are capable of inducing

Query Match 100.0%; Score 104; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIEAANSKD 20

CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KYNFPHGMLDLEETIAANSKD 20
DB 922 KYNFPHGMLDLEETIAANSKD 941
RESULT 5
AAW34074
ID AAW34074 standard; protein; 2227 AA.
XX
AC AAW34074;
XX
DT 17-OCT-2003 (revised)
DT 27-APR-1998 (first entry)
XX
DE /label= VP4 = 1A
XX
DE /label= VP2 = 1B
XX
KW /label= VP3 = 1C
XX
KW /label= VP1 = 1D
XX
OS Hepatitis A virus; strain HM-175.
XX
FH Key Location/Qualifiers
FT Protein 1..23
FT /label= VP4
FT Protein 24..245
FT /label= VP2
FT Protein 246..491
FT /label= VP3
FT Protein 492..791
FT /label= VP1
FT Protein 792..980
FT /label= 2A
FT Protein 981..1087
FT /label= 2B
FT Protein 1088..1422
FT /label= 2C
FT Protein 1423..1496
FT /label= 3A
FT Protein 1497..1519
FT /label= 3B
FT Protein 1520..1738
FT /label= 3C
FT Protein 1739..2227
FT /label= 3D
XX
PN WO9740166-A2.
XX
PI Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstone SM;
PI Daemer RJ, Gust ID;
XX
XX WPI; 1990-075557/10.
DR N-PSDB; AAQ03512.
XX
XX Vaccine against hepatitis A virus infection - comprises novel attenuated
PT hepatitis A virus strain.
XX
XX Claim 1; Fig 1; 10pp; English.
XX
XX The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-

DB 922 KYNFPHGMLDLEETIAANSKD 941
RESULT 4
AAR05697
ID AAR05697 standard; protein; 2227 AA.
XX
AC AAR05697;
XX
DT 24-OCT-2003 (revised)
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 15-AUG-1990 (first entry)
XX
XX Attenuated hepatitis A virus.
XX
XX Hepatitis A virus; vaccine; attenuated.
XX
XX Hepatitis A virus; strain HM-175.
XX
XX Key Location/Qualifiers
FT Region 1..23
FT /label= VP4 = 1A
FT Region 24..245
FT /label= VP2 = 1B
FT Region 246..491
FT /label= VP3 = 1C
FT Region 492..791
FT /label= VP1 = 1D
FT Region 792..980
FT /label= 2A
FT Region 981..1087
FT /label= 2B
FT Region 1088..1422
FT /label= 2C
FT Region 1423..1496
FT /label= 3A
FT Region 1497..1519
FT /label= 3B = VPg
FT Region 1520..1738
FT /label= 3C
FT Region 1739..2227
FT /label= 3D
XX
XX US4894228-A.
XX
XX 16-JAN-1990.
XX
XX 12-JUL-1988; 88US-00217824.
XX
XX 19-SEP-1984; 84US-00652067.
XX
XX 09-SEP-1986; 86US-00905146.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstone SM;
XX Daemer RJ, Gust ID;
XX
XX WPI; 1990-075557/10.
DR N-PSDB; AAQ03512.
XX
XX Vaccine against hepatitis A virus infection - comprises novel attenuated
PT hepatitis A virus strain.
XX
XX Claim 1; Fig 1; 10pp; English.
XX
XX The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-

PS Disclosure; Fig 13A-D; 66pp; English.

XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA construct (I) comprises a genome of HAV, where the genome is a human attenuated HAV genome in which a region of the 2C gene has been replaced by a corresponding region from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The region of the 2C gene from AGM-27 contained in the construct preferably encodes amino acids 120-328 of the 2C protein, amino acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (1) or its RNA transcript, can be used as a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can also be used to stimulate the production of protective antibodies in the mammal. (Updated on 17-OCT-2003 to standardise OS field)

XX

SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEIAANSKD 20
|||||
Db 922 KVNFPFGMLDLEIAANSKD 941

RESULT 6
AAB18609
ID AAB18609 standard; protein; 2227 AA.
XX
AC AAB18609;
XX
XX
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
XX
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW HAV 4380.
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
XX
PD 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-00475886.
XX
PR 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.
XX N-PSDB; AAA75478.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.
XX
XX Disclosure; Col 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A virus (HAV) of the invention, designated HAV 4380. The sequence is produced by modifying wild type HAV strain HM-174. The HAV of the invention are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans

XX

PS Disclosure; Fig 13A-D; 66pp; English.

XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA construct (I) comprises a genome of HAV, where the genome is a human attenuated HAV genome in which a region of the 2C gene has been replaced by a corresponding region from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The region of the 2C gene from AGM-27 contained in the construct preferably encodes amino acids 120-328 of the 2C protein, amino acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (1) or its RNA transcript, can be used as a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can also be used to stimulate the production of protective antibodies in the mammal. (Updated on 17-OCT-2003 to standardise OS field)

XX

SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEIAANSKD 20
|||||
Db 922 KVNFPFGMLDLEIAANSKD 941

RESULT 8
AAB18608
ID AAB18608 standard; protein; 2227 AA.
XX
AC AAB18608;

XX 15-JAN-2001 (first entry)
DT Amino acid sequence of passage 35 Hepatitis A virus called P-35.
DE HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW P-35 virus.
XX Hepatitis A virus.
OS US6113912-A.
XX 05-SEP-2000.
PD 07-JUN-1995; 95US-00475886.
XX 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.
DR N-PSDB; AAA75477.
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type.
XX Disclosure; Col 67-78; 72pp; English.
XX The present sequence is derived from passage 35 of a wild type hepatitis
CC A virus (HAV) strain HM-174. The resulting virus is designated P-35
CC virus. The sequence is modified to produce HAV which are adapted to
CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
CC as a live vaccine for prophylaxis of hepatitis A in humans and other
CC primates
XX
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNFPFGMLDLEEIAANSKD 20
DB 922 KVNFPFGMLDLEEIAANSKD 941
RESULT 9
AAE19899
ID AAE19899 standard; protein; 2227 AA.
XX
XX AC AAE19899;
XX
DT 18-JUN-2002 (first entry)
XX
DE Hepatitis A virus (HAV) protein.
XX
KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
OS Hepatitis A virus.
XX
XX WO200213855-A2.
XX
XX 21-FEB-2002.
PD
XX 15-AUG-2001; 2001WO-IB001808.
XX

PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
PA
XX Sallberg M, Hultgren C;
PI
XX WPI; 2002-241837/29.
DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Claim 11; Page 82-87; 120pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is hepatitis A virus (HAV) protein
XX
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNFPFGMLDLEEIAANSKD 20
DB 922 KVNFPFGMLDLEEIAANSKD 941
RESULT 10
ABG31729
ID ABG31729 standard; protein; 2227 AA.
XX
XX AC ABG31729;
XX
DT 29-AUG-2003 (revised)
DT 29-NOV-2002 (first entry)
XX
DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
KW HAV 4380.
XX
XX Hepatitis A virus; strain HM-175.
OS
XX US6423318-B1.
PN
XX 23-JUL-2002.
PD
XX 31-AUG-2000; 2000US-00653499.
XX
XX 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
PI
XX WPI; 2002-680946/73.
DR N-PSDB; ABS52789.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
PT

XX PS Disclosure; Col 93-104; 71pp; English.

XX CC The invention relates to a polynucleotide which encodes a hepatitis A

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

CC line). The polynucleotide is useful for preparing a vaccine against

CC hepatitis A virus infection. This sequence represents an attenuated

CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to

CC standardise OS field)

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 5; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.1e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLBEIAANSKD 20

DB 922 KVNPHGMLDLBEIAANSKD 941

RESULT 11

ABG31727

ID ABG31727 standard; protein; 2227 AA.

XX AC ABG31727;

XX DT 29-AUG-2003 (revised)

DT 29-NOV-2002 (first entry)

XX DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.

XX KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.

XX OS Hepatitis A virus; strain HM-175.

XX PN US6423318-B1.

XX PD 23-JUL-2002.

XX PE 31-AUG-2000; 2000US-00653499.

XX PR 17-SEP-1993; 93WO-US008610.

PR 17-APR-1995; 95US-00397232.

PR 07-JUN-1995; 95US-00475886.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

DR WPI; 2002-680946/73.

DR N-PSDB; ABS52787.

XX PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.

XX PS Disclosure; Fig 6; 71pp; English.

XX CC The invention relates to a polynucleotide which encodes a hepatitis A

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

CC line). The polynucleotide is useful for preparing a vaccine against

CC hepatitis A virus infection. This sequence represents a hepatitis A virus

CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS

CC field)

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 5; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.1e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLBEIAANSKD 20

DB 922 KVNPHGMLDLBEIAANSKD 941

RESULT 12

ABG31728

ID ABG31728 standard; protein; 2227 AA.

XX AC ABG31728;

XX DT 29-NOV-2002 (first entry)

XX DE Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.

XX KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;

XX virucide; mutant; PHAV/7; mutcin.

XX OS Hepatitis A virus; strain HM-175.

XX OS Synthetic.

XX PH Key Location/Qualifiers

FT Misc-difference 764 /note= "Wild-type Glu substituted by Val"

FT Misc-difference 821 /note= "Wild-type Asn substituted by Ser"

FT Misc-difference 963 /note= "Wild-type Lys substituted by Arg"

FT Misc-difference 1052 /note= "Wild-type Ala substituted by Val"

FT Misc-difference 1062 /note= "Wild-type Gly substituted by Ala"

FT Misc-difference 1118 /note= "Wild-type Lys substituted by Met"

FT Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"

FT Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"

FT Misc-difference 1277 /note= "Wild-type Val substituted by Ile"

FT Misc-difference 1500 /note= "Wild-type His substituted by Tyr"

FT Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"

FT Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"

XX PN US6423318-B1.

XX PD 23-JUL-2002.

XX PE 31-AUG-2000; 2000US-00653499.

XX PR 17-SEP-1993; 93WO-US008610.

PR 17-APR-1995; 95US-00397232.

PR 07-JUN-1995; 95US-00475886.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX WPI; 2002-680946/73.

DR N-PSDB; ABS52788.

XX PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.

XX PS Example 3; Col 67-78; 71pp; English.

XX CC The invention relates to a polynucleotide which encodes a hepatitis A

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

CC line). The polynucleotide is useful for preparing a vaccine against

CC hepatitis A virus infection. This sequence represents a hepatitis A virus

DE Wild type human hepatitis A virus strain HM-175.
 XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
 KW vaccine; MRC-5 cell; hepatitis infection.
 XX
 OS Hepatitis A virus; strain HM-175.
 XX
 PN US2002176859-A1.
 XX
 XX 28-NOV-2002.
 XX
 PF 29-APR-2002; 2002US-00135988.
 XX
 PR 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 PR 31-AUG-2000; 2000US-00653499.
 XX
 PA (FUNK/) FUNKHOUSER A W.
 PA (EMER/) EMERSON S U.
 PA (PURC/) PURCELL R H.
 PA (DHON/) D'HONDT E.
 XX
 PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 XX
 XX MPI; 2003-352605/02.
 DR N-PSDB; ABX93473.
 XX
 XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
 PT useful in vaccines for protecting primates against hepatitis infection
 PT and disease.
 XX
 PS Disclosure; Fig 6; 70pp; English.
 XX
 CC The invention describes a live hepatitis A virus (HAV) adapted to growth
 CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
 CC a vaccine for protecting primates against hepatitis infection and
 CC disease. This is the amino acid sequence of wild type human hepatitis A
 CC virus strain HM-175. (Updated on 23-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 6; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
 |||||
 Db 922 KVNPHGMLDLEIAANSKD 941

Search completed: March 15, 2004, 13:59:58
 Job time : 45.2353 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:15:47 ; Search time 11.2941 Seconds
(without alignments)
91.421 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104
Sequence: 1 KVNPPHGMLEIEIAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	2227	3	US-08-475-886-2
2	104	100.0	2227	3	US-08-475-886-4
3	104	100.0	2227	3	US-08-475-886-6
4	104	100.0	2227	3	US-08-397-232-2
5	104	100.0	2227	3	US-08-397-232-4
6	104	100.0	2227	3	US-09-171-387-2
7	104	100.0	2227	4	US-09-653-499-2
8	104	100.0	2227	4	US-09-653-499-4
9	104	100.0	2227	4	US-09-653-499-6
10	104	100.0	2227	4	US-10-104-966-12
11	104	100.0	2227	4	US-10-135-988-2
12	104	100.0	2227	4	US-10-135-988-4
13	104	100.0	2227	4	US-10-135-988-6
14	45	43.3	417	2	US-09-099-677A-6
15	45	43.3	417	3	US-09-261-471-6
16	44	42.3	432	2	US-09-099-677A-3
17	44	42.3	432	3	US-09-261-471-3
18	43	41.3	98	3	US-09-181-487-2
19	43	41.3	99	4	US-09-227-357-219
20	41.5	39.9	331	4	US-09-634-238-247
21	41	39.4	125	4	US-09-621-976-5826
22	41	39.4	311	4	US-09-134-000C-4383
23	40	38.5	589	4	US-09-939-309-8
24	40	38.5	589	4	US-09-849-180-8
25	40	38.5	589	4	US-09-356-643B-2
26	40	38.5	895	4	US-09-614-512-194
27	39	37.5	226	4	US-09-198-452A-210

Sequence 6363, Ap
Sequence 8698, Ap
Sequence 27913, A
Sequence 31, Appl
Sequence 31, Appl
Sequence 16, Appl
Sequence 19, Appl
Sequence 22, Appl
Sequence 25, Appl
Sequence 28, Appl
Sequence 31, Appl
Sequence 24, Appl
Sequence 90, Appl
Sequence 95, Appl
Sequence 92, Appl
Sequence 125, App
Sequence 6582, Ap
Sequence 3897, Ap

US-09-543-681A-6363
US-09-489-039A-8698
US-09-252-991A-27913
US-08-126-564A-31
PCT-US94-09143-31
US-09-115-475-16
US-09-115-475-19
US-09-115-475-22
US-09-115-475-25
US-09-115-475-28
US-09-115-475-31
US-09-004-838-24
US-09-004-838-90
US-09-004-838-95
US-09-004-838-92
US-09-543-681A-6582
US-09-134-001C-3897

ALIGNMENTS

RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE H
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPPHGMLEIEIAANSKD 20
|||
Db 922 KVNPPHGMLEIEIAANSKD 941

RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232

; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLBEIAANSKD 20
DB 922 KVNFPFGMLDLBEIAANSKD 941

RESULT 5
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLBEIAANSKD 20
DB 922 KVNFPFGMLDLBEIAANSKD 941

RESULT 6
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999

; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLBEIAANSKD 20
DB 922 KVNFPFGMLDLBEIAANSKD 941

RESULT 3
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLBEIAANSKD 20
DB 922 KVNFPFGMLDLBEIAANSKD 941

RESULT 4
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227

;; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feller
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPHGMLDLEIAANSKD 20
Db 922 KVNPHGMLDLEIAANSKD 941

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPHGMLDLEIAANSKD 20
Db 922 KVNPHGMLDLEIAANSKD 941

RESULT 8
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

;; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPHGMLDLEIAANSKD 20
Db 922 KVNPHGMLDLEIAANSKD 941

RESULT 9
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPHGMLDLEIAANSKD 20
Db 922 KVNPHGMLDLEIAANSKD 941

RESULT 10
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND

```

; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP 23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KVNFPHGMLDLLEIAANSKD 20
Db      922 KVNFPHGMLDLLEIAANSKD 941

RESULT 11
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KVNFPHGMLDLLEIAANSKD 20
Db      922 KVNFPHGMLDLLEIAANSKD 941

RESULT 12
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988

; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP 23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KVNFPHGMLDLLEIAANSKD 20
Db      922 KVNFPHGMLDLLEIAANSKD 941

RESULT 13
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match          100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KVNFPHGMLDLLEIAANSKD 20
Db      922 KVNFPHGMLDLLEIAANSKD 941

RESULT 14
US-09-099-677A-6
; Sequence 6, Application US/09099677A
; Patent No. 5965369
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN SUCCINYL-COENZYME A SYNTHETASE HOLOENZYME
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/099,677
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CERONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:23:07 ; Search time 23.1765 Seconds
(without alignments)
182.213 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNFPHGMLDERIANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	20	10	US-09-171-432A-44
2	104	100.0	2227	9	US-09-929-955-12
3	104	100.0	2227	13	US-10-104-966-12
4	104	100.0	2227	13	US-10-135-988-2
5	104	100.0	2227	13	US-10-135-988-4
6	104	100.0	2227	13	US-10-135-988-6
7	98	94.2	352	14	US-10-272-459-45
8	98	94.2	980	14	US-10-272-459-41
9	53	51.0	20	10	US-09-171-432A-45
10	46	44.2	609	15	US-10-369-493-21242
11	43	41.3	98	9	US-09-978-295A-616
12	43	41.3	98	9	US-09-978-697-616
13	43	41.3	98	9	US-09-978-192A-616
14	43	41.3	98	9	US-09-999-832A-616
15	43	41.3	98	10	US-09-978-189-616

16	43	41.3	98	10	US-09-978-608A-616	Sequence 616, App
17	43	41.3	98	10	US-09-978-585A-616	Sequence 616, App
18	43	41.3	98	10	US-09-978-191A-616	Sequence 616, App
19	43	41.3	98	10	US-09-978-403A-616	Sequence 616, App
20	43	41.3	98	10	US-09-978-564A-616	Sequence 616, App
21	43	41.3	98	10	US-09-999-833A-616	Sequence 616, App
22	43	41.3	98	10	US-09-981-915A-616	Sequence 616, App
23	43	41.3	98	10	US-09-978-824-616	Sequence 616, App
24	43	41.3	98	10	US-09-918-585A-616	Sequence 616, App
25	43	41.3	98	10	US-09-978-423A-616	Sequence 616, App
26	43	41.3	98	10	US-09-978-193A-616	Sequence 616, App
27	43	41.3	98	10	US-09-999-830A-616	Sequence 616, App
28	43	41.3	98	10	US-09-978-757A-616	Sequence 616, App
29	43	41.3	98	10	US-09-978-187B-616	Sequence 616, App
30	43	41.3	98	10	US-09-978-643A-616	Sequence 616, App
31	43	41.3	98	10	US-09-978-375A-616	Sequence 616, App
32	43	41.3	98	10	US-09-978-298A-616	Sequence 616, App
33	43	41.3	98	10	US-09-978-188A-616	Sequence 616, App
34	43	41.3	98	10	US-09-978-681A-616	Sequence 616, App
35	43	41.3	98	10	US-09-978-194A-616	Sequence 616, App
36	43	41.3	98	10	US-09-999-829A-616	Sequence 616, App
37	43	41.3	98	10	US-09-978-299A-616	Sequence 616, App
38	43	41.3	98	10	US-09-978-544A-616	Sequence 616, App
39	43	41.3	98	10	US-09-978-665A-616	Sequence 616, App
40	43	41.3	98	10	US-09-978-802A-616	Sequence 616, App
41	43	41.3	98	14	US-10-227-884-106	Sequence 106, App
42	43	41.3	98	14	US-10-230-163-106	Sequence 106, App
43	43	41.3	98	14	US-10-230-338-106	Sequence 106, App
44	43	41.3	98	14	US-10-218-631-106	Sequence 106, App
45	43	41.3	98	14	US-10-017-081A-616	Sequence 616, App

ALIGNMENTS

RESULT 1

US-09-171-432A-44
; Sequence 44, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1327
US-09-171-432a-44

Query Match      100.0%; Score 104; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPPHGMLEIEIAANSKD 20
Db 1 KVNPPHGMLEIEIAANSKD 20

RESULT 2
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      100.0%; Score 104; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPPHGMLEIEIAANSKD 20
Db 922 KVNPPHGMLEIEIAANSKD 941

RESULT 3
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29

Query Match      100.0%; Score 104; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPPHGMLEIEIAANSKD 20
Db 922 KVNPPHGMLEIEIAANSKD 941

RESULT 4
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match      100.0%; Score 104; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPPHGMLEIEIAANSKD 20
Db 922 KVNPPHGMLEIEIAANSKD 941

RESULT 5
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
```

```
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match      100.0%; Score 104; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KVNPHGMLDLEEAANSKD 20
DB      922 KVNPHGMLDLEEAANSKD 941

RESULT 6
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: FURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match      100.0%; Score 104; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KVNPHGMLDLEEAANSKD 20
DB      922 KVNPHGMLDLEEAANSKD 941

RESULT 7
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match      94.2%; Score 98; DB 14; Length 352;
Best Local Similarity 95.0%; Pred. No. 2.9e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KVNPHGMLDLEEAANSKD 20
DB      922 KVNPHGMLDLEEAANSKD 941

RESULT 8
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match      94.2%; Score 98; DB 14; Length 980;
Best Local Similarity 95.0%; Pred. No. 9.3e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KVNPHGMLDLEEAANSKD 20
DB      922 KVNPHGMLDLEEAANSKD 941

RESULT 9
US-09-171-432A-45
; Sequence 45, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
```



```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1328
US-09-171-432A-45

Query Match 51.0%; Score 53; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 DLEETAAASKD 20
Db 1 DLEETAAASKD 11

RESULT 10
US-10-369-493-21242
; Sequence 21242, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21242
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21242

Query Match 44.2%; Score 46; DB 15; Length 609;
Best Local Similarity 46.7%; Pred. No. 34;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 HGMLDLEETAAASKD 20
Db 274 HGQVDLSLGCANAE 288

RESULT 11
US-09-978-295A-616
; Sequence 616, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fotg, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
```

;
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414

;
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 9; Length 98;

Best Local Similarity 44.4%; Pred.No.13; Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPPHGMLEIEIAANSKD 20
DB 40 NLPEGVADLTQIDVNVQD 57

RESULT 12

US-09-978-697-616
; Sequence 616, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;

APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: Acids Encoding the Same
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 9; Length 98;
Best Local Similarity 44.4%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 NPPHGMLEETAAANSKD 20
Db 40 NLPEGVADLTQDVNVQD 57

RESULT 13

US-09-978-192A-616
Sequence 616, Application US/09978192A
Patent No. US20020177553A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08

;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627

;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085562
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 9; Length 98;
Best local Similarity 44.4%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NFPHGMLDLEEAASND 20

Db 40 NLPEGVADLTQIDNVQD 57

RESULT 14

US-09-959-832A-616
; Sequence 616, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C63
; CURRENT APPLICATION NUMBER: US/09/999,832A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30

PRIOR FILING DATE: 1998-04-09	PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23	PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27	PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28	PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05	PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-17	PRIOR APPLICATION NUMBER: 60/085338

OM protein - protein search, using sw model
Run on: March 15, 2004, 13:50:08 ; Search time 10 Seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KNFPHGMLDLBEIAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	104	100.0	2227	1 GNNYHM	genome polyprotein
2	104	100.0	2227	1 GNNYHR	genome polyprotein
3	104	100.0	2227	1 GNNYK	genome polyprotein
4	104	100.0	2227	1 GNNYHB	genome polyprotein
5	101	97.1	2230	1 GNNYSA	genome polyprotein
6	47	45.2	333	2 A12485	hypothetical prote
7	47	45.2	387	2 B86669	RecA protein [Impo
8	47	45.2	506	1 S85522	Glycine-tRNA ligas
9	47	45.2	1060	2 F88710	protein C01GS.4 [i
10	47	45.2	1079	2 T30996	hypothetical prote
11	46	44.2	259	2 F82087	deoxyribose-phosph
12	46	44.2	444	2 C9768	conserved hypotet
13	46	44.2	677	2 H64574	DNA topoisomerase
14	45	43.3	195	2 T08812	probable succinate
15	45	43.3	383	2 F96582	hypothetical prote
16	45	43.3	417	2 A44529	succinate-CoA liga
17	45	43.3	852	2 B72685	hypothetical prote
18	45	43.3	935	2 T19011	hypothetical prote
19	45	43.3	1213	2 T19835	hypothetical prote
20	44	42.3	223	2 T37962	hypothetical prote
21	44	42.3	344	2 D97761	hypothetical prote
22	44	42.3	500	2 S80508	ANP1 protein - yea
23	43.5	41.8	373	2 A93773	hypothetical prote
24	43	41.3	140	2 E81659	conserved hypotet
25	43	41.3	141	2 G71501	hypothetical prote
26	43	41.3	257	2 AF2592	2-deoxyribose-5-ph
27	43	41.3	259	2 H97374	hypothetical prote
28	43	41.3	309	2 H71089	hypothetical prote
29	43	41.3	327	2 T00876	probable glycerate

30 43 41.3 440 2 B71858
31 43 41.3 440 2 H64658
32 43 41.3 609 2 AB0955
33 43 41.3 917 2 T21870
34 42.5 40.9 599 2 A86810
35 42 40.4 222 2 C82343
36 42 40.4 229 2 S77449
37 42 40.4 259 2 D91296
38 42 40.4 259 2 F86137
39 42 40.4 290 2 S19426
40 42 40.4 309 2 T33259
41 42 40.4 342 2 C73113
42 42 40.4 365 2 B48945
43 42 40.4 384 2 E82088
44 42 40.4 495 2 D64578
45 42 40.4 539 2 G70520

adenylosuccinate 1
adenylosuccinate 1
glutamine-fructose
hypothetical prote
1-deoxyxylulose-5-
conserved hypotet
hydrogenase expres
2-deoxyribose-5-ph
2-deoxyribose-5-ph
hypothetical prote
hypothetical prote
hypothetical prote
recombination prot
conserved hypotet
conserved hypotet
probable esp prote

ALIGNMENTS

RESULT 1
GNNYHM
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C;Accession: A25981
R;Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A;Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with d
A;Reference number: A25981; MUID:87061253; PMID:3023706
A;Accession: A25981
A;Molecule type: Genomic RNA
A;Residues: 1-2227 <COR>
A;Cross-references: EMBL:M14707; NID:G329582; PIDN:AAA45465.1; PID:G329583
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltran
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-245/Product: coat protein 1B #status predicted <VP2>
F;246-491/Product: coat protein 1C #status predicted <VP3>
F;492-791/Product: coat protein 1D #status predicted <VP1>
F;792-980/Product: core protein 2A #status predicted <C2A>
F;981-1087/Product: core protein 2B #status predicted <C2B>
F;1088-1422/Product: core protein 2C #status predicted <C2C>
F;1423-1496/Product: protein 3A #status predicted <C3A>
F;1497-1519/Product: protein 3B #status predicted <C3B>
F;1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F;1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNFPHGMLDLBEIAANSKD 20
Db 922 KNFPHGMLDLBEIAANSKD 941

RESULT 2

GNNYHR
genome polyprotein - human hepatitis A virus
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pr
NA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C;Accession: A03903
R;Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nes
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A;Title: Primary structure and gene organization of human hepatitis A virus.

A/Reference number: A03903; MUID:85190549; PMID:2986127
A/Accession: A03903
A/Molecule type: Genomic RNA
A/Residues: 1-2227 <NAJ>
A/Cross-references: GB:K02990; NID:G329596; PIDN:AAA45472.1; PID:G329597
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: coat protein 2A #status predicted <C2A>
F:981-1076/Product: coat protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPPHGMGLDLERIAANSKD 20
|||||
Db 922 KVNPPHGMGLDLERIAANSKD 941

RESULT 3
GNMYMK
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C/Species: human hepatitis A virus
A/Note: host Homo sapiens (man)
C/Date: 30-Jun-1998 #sequence_revision 30-Jun-1998 #text_change 16-Jul-1999
R/Accession: A94149; A25914; A94508
R/Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daewer, R.J.; Feinstein, S.M.; Purcell, R.
Proc. Natl. Acad. Sci. U.S.A. 84: 2497-2501, 1987
A/Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with HAV
A/Reference number: A94149; MUID:87175701; PMID:3031866
A/Accession: A94149
A/Status: nucleic acid sequence not shown
A/Molecule type: Genomic RNA
A/Residues: 1-2227 <COH>
A/Cross-references: EMBL:M16632; NID:G329594; PIDN:AAA45471.1; PID:G329595
A/Note: submitted to GenBank, August 1987
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: coat protein 2A #status predicted <P2A>
F:981-1076/Product: coat protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPPHGMGLDLERIAANSKD 20
|||||
Db 922 KVNPPHGMGLDLERIAANSKD 941

RESULT 4
GNMYHB
genome polyprotein - human hepatitis A virus (strain MBB)
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 2D; core protein 3A; core protein 3B; core protein 3C; core protein 3D
Vg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C/Species: human hepatitis A virus
A/Note: host Homo sapiens (man)
C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C/Accession: J03003
R/Paul, A.V.; Tada, H.; von der Helm, K.; Missel, T.; Kiehn, R.; Wimmer, E.; Deinhardt, V.
Virus Res. 8, 153-171, 1987
A/Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolated from a patient with acute hepatitis A)
A/Reference number: J03003; MUID:88045071; PMID:2823500
A/Accession: J03003
A/Molecule type: Genomic RNA
A/Residues: 1-2227 <PAU>
A/Cross-references: EMBL:M20273
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrolase
F:1-23/Product: coat protein 1A #status predicted <VP1>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-431/Product: coat protein 1C #status predicted <VP3>
F:432-836/Product: coat protein 1D #status predicted <VP4>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: core protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPG>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPPHGMGLDLERIAANSKD 20
|||||
Db 922 KVNPPHGMGLDLERIAANSKD 941

RESULT 5
GNMYSA
genome polyprotein - simian hepatitis A virus (strain AGM-27)
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 2D; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C/Species: simian hepatitis A virus
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C/Accession: A30470; S04885; S03965
R/Tsarev, S.A.
submitted to JIPID, April 1991
A/Reference number: A30470
A/Accession: A30470
A/Molecule type: Genomic RNA
A/Residues: 1-2230 <TSA>
A/Cross-references: GB:D00924; NID:G222597; PIDN:BA00766.1; PID:G222598
R/Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A/Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure with HAV
A/Reference number: J01080; MUID:91311420; PMID:1649901
A/Contents: annotation
A/Note: neither amino acid nor nucleotide sequence is given
R/Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik, V.
Submitted to the EMBL Data Library, May 1989
A/Reference number: S04885
A/Accession: S04885
A/Molecule type: Genomic RNA
A/Residues: 1950-2164 <BAL>
A/Cross-references: EMBL:X15461; NID:G61971; PIDN:CAA33490.1; PID:9930268
R/Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik, V.
FEBS Lett. 247, 425-428, 1989
A/Title: Variations in genome fragments coding for RNA polymerase in human and simian hepatitis A virus
A/Reference number: S03965; MUID:89232168; PMID:2541023
A/Accession: S03965
A/Molecule type: Genomic RNA
A/Residues: 1960-2164 <BAL2>
A/Cross-references: EMBL:X15461
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; polyprotein
F:1-27/Product: coat protein 1A #status predicted <C1A>

F:28-249/Product: coat protein 1B #status predicted <CLB>
F:250-495/Product: coat protein 1C #status predicted <CLC>
F:496-795/Product: coat protein 1D #status predicted <CLD>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: core protein 3A #status predicted <P3A>
F:1499-1521/Product: core protein 3B #status predicted <P3B>
F:1522-1741/Product: core protein 3C #status predicted <P3C>
F:1742-2230/Product: core protein 3D #status predicted <P3D>

Query Match 97.1%; Score 101; DB 1; Length 2230;
Best Local Similarity 95.0%; Pred. No. 4.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEEAANSKD 20
DB 926 KVNFPFGMLDLEEAANSKD 945

RESULT 6
AI2485
hypothetical protein all7065 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AI2485
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, K.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2485
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <KUR>
A:Cross-references: GB:BA000020; PIDN:BA078149.1; PID:g17135603; GSPDB:GN00180
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all7065
A:Genome: plasmid

Query Match 45.2%; Score 47; DB 2; Length 333;
Best Local Similarity 40.0%; Pred. No. 5.6;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEEAANSKD 20
DB 57 KSGYPKGITLQEVVANVD 76

RESULT 7
B8669
RecA protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: B8669
R:Bolet, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, S. Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis subsp. lactis
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: B8669
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <STO>
A:Cross-references: GB:AB005176; PID:g12723223; PIDN:AAK04452.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA

Query Match 45.2%; Score 47; DB 2; Length 387;
Best Local Similarity 45.0%; Pred. No. 6.7;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEEAANSKD 20
DB 340 KVRTAAGLLDEAEVAETTED 359

RESULT 8
S58522
glycyl-tRNA ligase (EC 6.1.1.14) - Thermus aquaticus
N:Alternate names: glycyl-tRNA synthetase
C:Species: Thermus aquaticus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S58522
R:Logan, D.T.; Mazauric, M.H.; Kern, D.; Moras, D. EMBO J. 14, 4156-4167, 1995
A:Title: Crystal structure of glycyl-tRNA synthetase from Thermus thermophilus.
A:Reference number: S58522; MUID:96016187; PMID:7556056
A:Accession: S58522
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-506 <LOG>
A:Experimental source: strain HB8
A:Note: The source is designated as Thermus thermophilus
C:Superfamily: Mycoplasma genitalium glycine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 45.2%; Score 47; DB 1; Length 506;
Best Local Similarity 75.0%; Pred. No. 9.2;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 FPHGMLDLEEA 15
DB 299 FPHGSLELEGIA 310

RESULT 9
F88710
protein C01G5.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: F88710
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: AF5000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
A:Accession: F88710
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1060 <STO>
A:Cross-references: GB:chr_IV; PIDN:AB37736.1; PID:g1208856; GSPDB:GN00022; CESP:C01G5.4
C:Genetics:
A:Gene: C01G5.4
A:Map position: 4

Query Match 45.2%; Score 47; DB 2; Length 1060;
Best Local Similarity 36.8%; Pred. No. 22;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLEEAANSKD 20
DB 391 IGFSRGILDLKHVAGDARD 409

RESULT 10
T30996
hypothetical protein C01G5.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 29-Oct-1999
C:Accession: T30996
R:Bradshaw, H.; Stellyes, L.

submitted to the EMBL Data Library, August 1999
A:Description: The sequence of *C. elegans* cosmid C01G5.
A:Reference number: Z20956
A:Accession: T30996
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1079 <BRA>
A:Cross-references: EMBL:U50068; PIDN:AAB37736.2
A:Experimental source: strain Bristol N2
C:Genetics:
A:Map position: IV
A:Introns: 10/1; 31/1; 75/1; 108/1; 144/1; 927/2; 1056/2
A:Note: C01G5.4

Query Match 45.2%; Score 47; DB 2; Length 1079;
Best Local Similarity 36.8%; Pred. No. 22;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLERIAANSKD 20
DB 410 IGFSRGILDKHVAGDARD 428
: ||| : : : :
: ||| : : : :
: ||| : : : :

RESULT 11
F82087
deoxyribose-phosphate aldolase VC2350 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: F82087
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: F82087
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <HEI>
A:Cross-references: GB:AE004305; GB:AE003852; NID:g9556912; PIDN:AAF95493.1; GSFDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:

A:Gene: VC2350

A:Map position: 1
C:Superfamily: deoxyribose-phosphate aldolase

Query Match 44.2%; Score 46; DB 2; Length 259;
Best Local Similarity 52.9%; Pred. No. 6.2;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 NFPHGMLDLERIAANSK 19

DB 75 NFPHGNDDBIAVAETK 91
: ||| : : : :
: ||| : : : :
: ||| : : : :

RESULT 12
C89768
conserved hypothetical protein SA0083 [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89768
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogudma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89768
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <KUR>
A:Cross-references: GB:BA000018; PID:g13700003; PIDN:BAB41302.1; GSFDB:GN00149
A:Experimental source: strain N315

C:Genetics:
A:Gene: SA0083
Query Match 44.2%; Score 46; DB 2; Length 444;
Best Local Similarity 52.6%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLERIAANSKD 20
DB 384 VNIPHGKILNENIPFNKD 402
: ||| : : : :
: ||| : : : :
: ||| : : : :

RESULT 13
H64574
DNA topoisomerase I - *Helicobacter pylori* (strain 26695)
C:Species: *Helicobacter pylori*
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: H64574
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.F.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, I. Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64574
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-677 <TOM>
A:Cross-references: GB:AE000559; GB:AE000511; NID:g2313536; PIDN:AAD07502.1; PID:g23135
C:Superfamily: DNA topoisomerase I

Query Match 44.2%; Score 46; DB 2; Length 677;
Best Local Similarity 31.6%; Pred. No. 19;
Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLERIAANSK 19
DB 346 RITHPHALKOLEKVCSDAK 364
: ||| : : : :
: ||| : : : :
: ||| : : : :

RESULT 14
T08812
probable succinate-CoA ligase (GDP-forming) (EC 6.2.1.4) beta chain - human (fragment)
N:Alternate names: protein DKFZp586M2023.1; succinyl-CoA synthetase (GDP-forming) beta chain
C:Species: *Homo sapiens* (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C:Accession: T08812
R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A:Reference number: 216472
A:Accession: T08812
A:Molecule type: mRNA
A:Residues: 1-195 <ANS>
A:Cross-references: EMBL:AL050226
A:Experimental source: adult uterus; clone DKFZp586M2023
C:Genetics:

A:Note: DKFZp586M2023.1
C:Function:
A:Description: catalyzes the formation of succinyl-CoA from succinate with concomitant
A:Superfamily: succinate-CoA ligase (ADP-forming) beta chain
C:Keywords: acid-thiol ligase; coenzyme A; mitochondrion

Query Match 43.3%; Score 45; DB 2; Length 195;
Best Local Similarity 42.1%; Pred. No. 6.6;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLERIAANSKD 20
DB 65 VGSPPQGGVDIEVAASNP 83
: ||| : : : :
: ||| : : : :
: ||| : : : :

RESULT 15

F96582
 Hypothetical protein F511.25 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: F96582
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: F96582
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-383 <STO>
 A;Cross-references: GB:AE005173; NID:94587558; PIDN:AAD25789.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F511.25
 A;Map position: 1

Query Match 43.3%; Score 45; DB 2; Length 383;
 Best Local Similarity 44.4%; Pred. No. 14;
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NFPHGMLDLEIFIAANSKD 20
 :||:||||:
 Db 209 DYPVGLLDIEKILPPGKD 226

Search completed: March 15, 2004, 14:06:59
 Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:26:12 ; Search time 5.64706 Seconds
(without alignments)
184.415 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNPPHGMLEIRANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	2226	1 POLG_HPAV4	P26581 hepatitis a
2	104	100.0	2226	1 POLG_HPAV8	P26582 hepatitis a
3	104	100.0	2227	1 POLG_HPAVH	P08617 hepatitis a
4	104	100.0	2227	1 POLG_HPAVL	P06441 hepatitis a
5	104	100.0	2227	1 POLG_HPAVM	P13901 hepatitis a
6	101	97.1	2230	1 POLG_HPAVS	P14553 simian hepa
7	100	96.2	2226	1 POLG_HPAV2	P26580 hepatitis a
8	47	45.2	387	1 REC2_LACIA	Q01840 lactococcus
9	47	45.2	404	1 SCB2_MOUSE	Q2218 mus musculu
10	47	45.2	505	1 SYG_THERH	P56206 thermus the
11	46	44.2	258	1 DROC_VIBVY	Q7mi38 vibrio vuln
12	46	44.2	259	1 DROC_VIBCH	Q8xp17 vibrio chol
13	45	43.3	432	1 SCB2_HUMAN	Q96199 homo sapien
14	45	43.3	433	1 SCB2_PIG	P33590 sus scrofa
15	44	42.3	500	1 ANP1_YEAST	P32629 saccharomyc
16	43	41.3	98	1 SRG1_HUMAN	O75711 homo sapien
17	43	41.3	258	1 DROC_VIBPA	Q87m22 vibrio para
18	43	41.3	259	1 DROC_AGRTS	Q8uJ09 agrobacteri
19	43	41.3	357	1 UPB2_CHICK	Q57429 gallus gall
20	43	41.3	440	1 PUR8_HELPJ	Q32ka2 helicobacte
21	43	41.3	440	1 PUR8_HELPJ	P56468 helicobacte
22	43	41.3	608	1 GLMS_SALTI	Q8z2Q2 s glucosami
23	42	40.4	214	1 GLP1_PHANI	P45853 pharbitis n
24	42	40.4	259	1 DROC_ECO57	Q8xb36 escherichia
25	42	40.4	259	1 DROC_SHIFL	Q3ip02 shigella fl
26	42	40.4	290	1 YQ6E_YEAST	P25617 saccharomyc
27	41.5	39.9	406	1 CALB_PROSL	Q8gb19 proteus sp.
28	41.5	39.9	545	1 SGLP_DRONE	Q97y2 drosophila
29	41	39.4	193	1 YB7I_HABIN	P44339 haemophilus
30	41	39.4	218	1 Y363_RICPR	Q9zdg6 rickettsia
31	41	39.4	400	1 NIFS_ANAAZ	Q43884 anabaena az
32	41	39.4	400	1 NIFS_ANASP	P12623 anabaena sp
33	41	39.4	502	1 NU2C_MESVI	Q9muq6 mesostigma

ALIGNMENTS

RESULT 1

ID	POLG_HPAV4	STANDARD;	PRT;	2226 AA.
AC	P26581;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]			
OS	Hepatitis A virus (strain 43c).			
OC	Viruses; serNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC	Hepatovirus.			
OX	NCBI_TaxID=12095;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91162758; PubMed=1705995;			
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,			
RA	Croneans T., Jansen R.W.;			
RT	"Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."			
RL	J. Virol. 65:2056-2065 (1991).			
CC	-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).			
CC	-!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.			
CC	-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.			
CC	-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; M59809; AAA45469.1; ..			
DR	HEROPS; C03.005; ..			
DR	InterPro; IPR004004; Calici_pol_hel.			
DR	InterPro; IPR009003; Cys_Ser_trypsin.			
DR	InterPro; IPR006050; RNA_helicase.			
DR	InterPro; IPR007095; RNA_pol_DS_PS.			
DR	InterPro; IPR001205; RNA_pol_P3D.			
DR	InterPro; IPR007094; RNA_pol_P5vir.			
DR	InterPro; IPR008975; Viral_cap_coat.			
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.			
DR	Pfam; PF00910; RNA_helicase; 1.			
DR	PRINTS; PR00918; CALICIVIRUSN.			
DR	Polyprotein; Coat protein; Core protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.			
CHAIN	1 23			
FT	CHAIN 24 245			
FT	CHAIN 246 491			

Q8xmp3 clostridium
P17798 agrobacteri
O05973 rickettsia
P20504 vaccinia vi
P33053 variola vir
P07392 vaccinia vi
Q9p9t2 xylella fas
Q8dbt2 vibrio vuln
P46905 bacillus su
O88623 mus musculu
Q9jtcx0 neisseria m
P47493 mycoplasma

34 41 39.4 552 1 SYG_CLOPE
35 40.5 38.9 237 1 VIB8_AGRTS
36 40.5 38.9 434 1 UD8_RICPR
37 40.5 38.9 1286 1 RPOI_VACCC
38 40.5 38.9 1286 1 RPOI_VACCC
39 40.5 38.9 1287 1 RPOI_VACCV
40 40 38.5 126 1 YF81_XYLFA
41 40 38.5 258 1 DROC_VIBVU
42 40 38.5 310 1 YCKK_BACSU
43 40 38.5 353 1 UBF2_MOUSE
44 40 38.5 404 1 ISCS_NEIMA
45 40 38.5 446 1 SYG_MYCGE

```
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D) .
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA, 251107 MW, 403B4CA90B09BF75 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNFPHGMLDLEIAANSKD 20
DB 922 KNFPHGMLDLEIAANSKD 941

RESULT 2
POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jensen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59808; AAA5467.1; --
DR PDB; 1QAV; 15-MAY-00.
DR MEROPS; C03.005; -.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_ser_typsln.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR InterPro; IPR001205; RNA_pol_DS_Ps.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00650; RNA_dep_RNA_pol_1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
```

```
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA) .
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB) .
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC) .
FT CHAIN 492 794 COAT PROTEIN VP1 (PID) .
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA, 251292 MW, 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNFPHGMLDLEIAANSKD 20
DB 922 KNFPHGMLDLEIAANSKD 941

RESULT 3
POLG_HPAV8 STANDARD; PRT; 2227 AA.
AC P05617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REP.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
```

CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14114; AAA45475.1; -
CC EMBL; M14707; AAA45465.1; -
CC EMBL; M14707; AAA45466.1; ALT_INIT.
CC EMBL; M16632; AAA45471.1; -
CC PIR; A03905; A03905.
CC PIR; A25981; GNNYHM.
CC PIR; A94149; GNNYMK.
CC PDB; 1HAV; 23-DEC-96.
CC MEROPS; C03.005; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT VARIANT 77 77
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277
FT VARIANT 1500 1500
FT VARIANT 1805 1805
FT VARIANT 1930 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7A5B740A6 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVFPHGMGLDLEIAANSKD 20
Db 922 KVFPHGMGLDLEIAANSKD 941

RESULT 4
POLG HPVAVL STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain LA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Marryweather J., van Nest G., Dina D.;
RT "Primary structure and gene organization of human hepatitis A virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02990; AAA45472.1; -
CC PIR; A03903; GNNYHR.
CC MEROPS; C03.005; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVFPHGMGLDLEIAANSKD 20
Db 922 KVFPHGMGLDLEIAANSKD 941

RESULT 5
POLG HPVAVL STANDARD; PRT; 2227 AA.
ID POLG HPVAVL STANDARD; PRT; 2227 AA.
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain MBB).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=28233500;
 RA Paul A.V., Tada H., der Helm K., Wessel T., Kiehn R., Wimmer E.,
 RA Deinhardt F.;
 RT "The entire nucleotide sequence of the genome of human hepatitis A
 RT virus (isolate MBB)."
 RL Virus Res. 8:153-171(1987).
 CC [1-] CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
 CC (RNA)(N).
 CC [1-] SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC [1-] PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC [1-] MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
 CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC [1-] SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M20273; AAA5474.1; --
 DR MEROPS; C03.005; --
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral cap.coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol_1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUS.
 DR Polyprotein; Coat protein; Core protein; Transferase;
 DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
 FT CHAIN 837 980 CORE PROTEIN P2A.
 FT CHAIN 981 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
 FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
 Query Match 100.0%; Score 104; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 3.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KNVPHGMLDLIEIAANSKD 20
 |||||
 Db 922 KNVPHGMLDLIEIAANSKD 941
 |||||
 RESULT 6
 POLG HPVAs
 ID POLG_HPVA STANDARD; PRT; 2230 AA.

AC P14553;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Simian hepatitis A virus (strain AGM-27).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91311420; PubMed=1649901;
 RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
 RA Purcell R.H.;
 RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
 RT structure and growth in cell culture with other HAV strains."
 RL J. Gen. Virol. 72:1677-1683(1991).
 CC [2]
 CC SEQUENCE OF 1750-2164 FROM N.A.
 RX MEDLINE=89232168; PubMed=2541023;
 RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
 RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
 RT "Variations in genome fragments coding for RNA polymerase in human
 RT and simian hepatitis A viruses."
 RL FEBS Lett. 247:425-428(1989).
 CC [1-] CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
 CC (RNA)(N).
 CC [1-] SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC [1-] PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC [1-] SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D00924; BAA00766.1; --
 DR EMBL; X15461; CAA33490.1; --
 DR PIR; A30470; GNNYSA.
 DR MEROPS; C03.005; --
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral cap.coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol_1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUS.
 DR Polyprotein; Coat protein; Core protein; Transferase;
 DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 27 COAT PROTEIN VP4 (P1A).
 FT CHAIN 28 249 COAT PROTEIN VP2 (P1B).
 FT CHAIN 250 495 COAT PROTEIN VP3 (P1C).
 FT CHAIN 496 795 COAT PROTEIN VP1 (P1D).
 FT CHAIN 796 984 CORE PROTEIN P2A.
 FT CHAIN 985 1091 CORE PROTEIN P2B.
 FT CHAIN 1092 1426 CORE PROTEIN P2C.
 FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
 FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
 FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
 FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
 Query Match 97.1%; Score 101; DB 1; Length 2230;

Best Local Similarity 95.0%; Pred. No. 1.1e-08; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPPHGMLEIEIAANSKD 20
DB 926 KVNPPHGMLEIEIAANSKD 945
RESULT 7
POLG HPV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
RA Cremons T., Jansen R.W.,
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RL J. Virol. 65:2056-2065(1991).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; M59810; AAA5468.1; -
DR MEROPS; C03.005; -
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006005; RNA helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
FT CHAIN 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;
SQ SEQUENCE

Query Match 96.2%; Score 100; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPPHGMLEIEIAANSKD 20
DB 922 KVNPPHGMLEIEIAANSKE 941
RESULT 8
REC2 LACIA STANDARD; PRT; 387 AA.
ID REC2 LACIA STANDARD;
AC Q01840; Q9C1K3;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RecA protein, chromosomal (Recombinase A).
DE RecA OR L10354.
GN RecA
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ML3;
RX MEDLINE=92384590; PubMed=1514816;
RA Duwat P., Ehrlich S.D., Gruss A.;
RT "Use of degenerate primers for polymerase chain reaction cloning and
RT sequencing of the Lactococcus lactis subsp. lactis recA gene."
RL Appl. Environ. Microbiol. 58:2674-2678(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarre K.,
RA Weissbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; M88106; AAA25216.1; -
DR EMBL; AS006272; AAK04452.1; -
DR PIR; B48945; B48945.
DR PIR; B86669; B86669.
DR HSSP; P26345; 1G19.
DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001533; RecA.
DR Pfam; PF00154; recA; 1.
DR PRINTS; PR00142; RECA.
DR ProDom; PD000229; RecA; 1.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS00162; RECA_2; 1.
DR PROSITE; PS00163; RECA_3; 1.
DR DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding;
Complete proteome.
NP_BIND 80 87 ATP (BY SIMILARITY).

FT CONFLICT 288 288 R -> S (IN REF. 1).
 FT CONFLICT 298 298 D -> E (IN REF. 1).
 FT CONFLICT 332 333 EI -> DV (IN REF. 1).
 FT CONFLICT 343 343 T -> A (IN REF. 1).
 FT CONFLICT 350 351 EA -> DS (IN REF. 1).
 FT CONFLICT 357 387 TETSTKATKATKKEKVETEEIELEED -> BEETTA
 FT CONFLICT 387 387 FKN (IN REF. 1).
 SQ SEQUENCE 387 AA; 41477 MW; A8CAFC0BB27BF14F CRC64;
 Query Match 45.28; Score 47; DB 1; Length 387;
 Best Local Similarity 45.08; Pred. No. 2.7;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 1 KVFPHGMGLDEEIAANSKD 20
 DB 340 KVRTAHGLDEAEVAETED 359
 RESULT 9
 ID SCB2 MOUSE STANDARD; PRT; 404 AA.
 AC Q922I8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Succinyl-CoA ligase (GDP-forming) beta-chain, mitochondrial precursor
 DE Succinyl-CoA ligase (GDP-forming) beta-chain (SCS-betaG) (GTP-
 DE specific succinyl-CoA synthetase, beta subunit) (Fragment).
 GN SUCLG2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98438536; PubMed=9765291;
 RA Johnson J.D., Mahus J.G., Teva K., Milavetz B.I., Lambeth D.O.;
 RT "Genetic evidence for the expression of Atp- and Gtp-specific
 RT succinyl-CoA synthetases in multicellular eucaryotes.";
 RL J. Biol. Chem. 273:27580-27586(1998).
 CC -1- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA +
 CC phosphate.
 CC -1- PATHWAY: Substrate level phosphorylation step of the tricarboxylic
 CC acid cycle.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
 CC subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF058956; AAC64399.1; --
 CC HSRP; P07460; I8CU.
 CC MGD; MGI:1306824; Sncg2.
 CC InterPro; IPR003135; ATP-grasp.
 CC InterPro; IPR005809; CoA_lig_beta.
 CC InterPro; IPR005811; CoA_ligase.
 CC Pfam; PF02222; ATP-grasp/1.
 CC Pfam; PF00549; ligase-CoA; 1.
 CC TIGRFAMs; TIGR01016; succCoabeta; 1.
 CC PROSITE; PS01217; SUCCINYL_COA_LIG_3; 1.
 KW Ligase; Glycolysis; Tricarboxylic acid cycle; Mitochondrion;
 KW Transit peptide.
 FT NON TER 1
 FT TRANSIT <1 9 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 10 404 SUCCINYL-COA LIGASE [GDP-FORMING] BETA-

FT CONFLICT 404 AA; 43857 MW; 511757A75883B8C4 CRC64;
 SQ SEQUENCE 404 AA; 43857 MW; 511757A75883B8C4 CRC64;
 Query Match 45.28; Score 47; DB 1; Length 404;
 Best Local Similarity 47.48; Pred. No. 2.9;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 2 VNFPHGMGLDEEIAANSKD 20
 DB 140 VGSFQGRSDIEEVAASSPE 158
 RESULT 10
 ID SYG THETH STANDARD; PRT; 505 AA.
 AC P56206; O50551;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine--tRNA ligase) (GlyRS).
 GN GLYS.
 OS Thermus thermophilus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermaceae;
 OC Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.75 ANGSTROMS).
 RC STRAIN=HB8 / ATCC 27634;
 RX MEDLINE=96016187; PubMed=7556056;
 RA Logan D.T., Mazauric M.-H., Kern D., Moras D.;
 RT "Crystal structure of glycyl-tRNA synthetase from Thermus
 RT thermophilus.";
 RL EMBO J. 14:4156-4167(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB8 / ATCC 27634;
 RX MEDLINE=98149692; PubMed=9490048;
 RA Mazauric M.-H., Keith G., Logan D., Kreutzer R., Giege R., Kern D.;
 RT "Glycyl-tRNA synthetase from Thermus thermophilus -- wide structural
 RT divergence with other prokaryotic glycyl-tRNA synthetases and
 RT functional inter-relation with prokaryotic and eukaryotic glycylation
 RT systems.";
 RL Eur. J. Biochem. 251:744-757(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) = AMP + diphosphate
 CC + glycyl-tRNA(Gly).
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ222643; CAA10903.1; --
 CC PDB; 1ATI; 07-JUL-97.
 CC PDB; 1B76; 28-JAN-99.
 CC PDB; 1GGW; 28-JAN-99.
 CC HAVAP; MF 00253; -; 1.
 CC InterPro; IPR004154; HGTP-anticodon.
 CC InterPro; IPR002314; tRNA-synt_2b.
 CC InterPro; IPR002315; tRNA-synt_gly.
 CC InterPro; IPR006195; tRNA_ligase_II.
 CC Pfam; PF03129; HGTP-anticodon; 1.
 CC Pfam; PF00587; tRNA-synt_2b; 1.
 CC PRINTS; PR01043; TRNASYNTHGLY.
 CC TIGRFAMs; TIGR00389; GLYS dimeric; 1.
 CC PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW 3D-structure.

[illegible]

[illegible]

RP SEQUENCE FROM N.A.
RC TISSUE=Placenta; and Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausser R.I., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner I., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh P.,
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.N., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Mazny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schultz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schmerch A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 29-432 FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=98438336; PubMed=9765291;
RC Johnson J.D., Mehru J.G., Tewks K., Milavet B.I., Lambeth D.O.;
RT "Genetic evidence for the expression of ATP- and GTP-specific
RT succinyl-CoA synthetases in multicellular eucaryotes";
R] J. Biol. Chem. 273:27580-27586(1998).
RN [3]
RP SEQUENCE OF 234-432 FROM N.A.
RA Mei Q., Yu W., Gibbs R.A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.

-I- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA + phosphate.
CC -I- PATHWAY: Substrate level phosphorylation step of the tricarboxylic acid cycle.
CC CC
CC -I- SUBUNIT: Tetradimer of an alpha and a beta chain.
CC CC
CC -I- SUBCELLULAR LOCATION: Mitochondrial.
CC CC
CC -I- SIMILARITY: Belongs to the succinate/malate CoA ligase beta subunit family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licensel@isb-sib.ch).

EMBL; BC007716; AAH07716.1; --
DR EMEL; BC019868; AAH19868.1; --
DR EMEL; AF058954; AAC64397.1; --
DR EMEL; AF131748; AAD20032.1; --
DR HSPP; P07460; 1SCU.
DR Genew; HGNC:11450; SUCLG2.
DR MTM; 603922; --
DR GO; GO:0004776; F:succinate-CoA ligase (GDP-forming) activity; NAS.
DR GO; GO:0006104; P:succinyl-CoA metabolism; NAS.
DR InterPro; IPRO03135; ATP-grasp.
DR InterPro; IPRO05809; CoA_lig_beta.
DR InterPro; IPRO05811; CoA_ligase.
DR Pfam; PF02222; ATP-grasp; 1.
DR Pfam; PF00549; ligase-coa; 1.
DR TRIGRAMS; TIGR01016; sucCoAbeta; 1.
DR PROSITE; PS01217; SUCCINYL COA LIg_3; 1.
KW Ligase; Glycocalysis; Tricarboxylic acid cycle; Mitochondrion;
KM Transit peptide.
FT NON_TER 1
FT TRANSIT <1 37 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 38 432 SUCCINYL-COA LIGASE [GDP-FORMING] BETA-

FT FT CONFLICT 174 175 CHAIN.
 SQ SEQUENCE 432 AA; 46535 MW; EB4D025B42EA7BE0 CRC64;
 Query Match 43.3%; Score 45; DB 1; Length 432;
 Best Local Similarity 42.1%; Pred. No. 6.7;
 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLBIEIAANSKD 20
 Db 168 VGSPQGGVDIEVAASNP 186

RESULT 14
 SCB2_PIG
 ID SCB2_PIG STANDARD; PRT; 433 AA.
 AC P53590; Q95279;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Succinyl-CoA ligase [GDP-forming] beta-chain, mitochondrial precursor
 DE (EC 6.2.1.4) (Succinyl-CoA synthetase, beta chain) (SCS-beta) (GTP-
 specific succinyl-CoA synthetase beta subunit) (Fragment).
 GN SUCLG2.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-41.
 RC TISSUE=Heart;
 RX MEDLINE=94004462; PubMed=8401211;
 RA Bailey D.L., Wolodko W.T., Bridger W.A.;
 RT "Cloning, characterization, and expression of the beta subunit of pig
 RT heart succinyl-CoA synthetase";
 RL Protein Sci. 2:1255-1262(1993).
 RN [2]
 RP SEQUENCE OF 1-82 FROM N.A.
 RC TISSUE=Small intestine;
 RX MEDLINE=96327607; PubMed=8672129;
 RA Winteroe A.K., Fredholm M., Davies W.;
 RT "Evaluation and characterization of a porcine small intestine cDNA
 RT library: analysis of 899 clones";
 RL Mamm. Genome 7:509-517(1996).
 CC -!- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA +
 CC phosphate.
 CC -!- PATHWAY: Substrate level phosphorylation step of the tricarboxylic
 CC acid cycle.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.
 CC -!- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
 CC subunit family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L06944; AA31120.1; ALT_INIT.
 CC EMBL; 261187; CAB03559.1; -;
 CC PDB; 1EUC; 27-JUL-00.
 CC PDB; 1EUD; 27-JUL-00.
 CC InterPro; IPR003135; ATP-grasp.
 CC InterPro; IPR005809; CoA_lig_beta.
 CC Pfam; PF02222; ATP-grasp; 1.
 CC Pfam; PF00549; ligase-CoA; 1.
 CC TIGRFAMs; TIGR01016; succCoABeta; 1.
 CC PROSITE; PS01217; SUCCINYL COA LIG 3; 1.
 KW Ligase; Glycolysis; Tricarboxylic acid cycle; Mitochondrion;

KW Transit peptide; 3D-structure.
 FT NON_TER 1
 FT TRANSIT <1 38 MITOCHONDRION.
 FT CHAIN 39 433 SUCCINYL-COA LIGASE [GDP-FORMING] BETA-
 FT CHAIN.
 SQ SEQUENCE 433 AA; 46803 MW; AA04B72BC1B80E24 CRC64;
 Query Match 43.3%; Score 45; DB 1; Length 433;
 Best Local Similarity 42.1%; Pred. No. 6.7;
 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLBIEIAANSKD 20
 Db 169 VGSPQGGVDIEVAASNP 187

RESULT 15
 ANP1_YEAST
 ID ANP1_YEAST STANDARD; PRT; 500 AA.
 AC P32629;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Mannan polymerase II complex ANP1 subunit (M-Pol II subunit ANP1)
 DE (Aminonitrophenyl propionediol resistance protein).
 GN ANP1 OR GEM3 OR YEL036C OR SYGP-ORF28.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-6441;
 RX MEDLINE=94016558; PubMed=8411151;
 RA Welnick L., Sherman F.;
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
 RT of Saccharomyces cerevisiae share a common ancestry";
 RL J. Mol. Biol. 233:372-388(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313264; PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 RA Huntcke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Borstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
 RL Nature 387:78-81(1997).
 RN [3]
 RP SEQUENCE OF 1-10, SUBCELLULAR LOCATION, AND SUBUNITS.
 RX MEDLINE=98096381; PubMed=9434768;
 RA Hashimoto H., Yoda K.;
 RT "Novel membrane protein complexes for protein glycosylation in the
 RT yeast Golgi apparatus";
 RL Biochem. Biophys. Res. Commun. 241:682-686(1997).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=95045392; PubMed=7957057;
 RA Chapman R.E., Munro S.;
 RT "The functioning of the yeast Golgi apparatus requires an ER protein
 RT encoded by ANP1, a member of a new family of genes affecting the
 RT secretory pathway";
 RL EMBO J. 13:4896-4907(1994).
 RN [5]
 RP ACTIVITY OF M-POL II COMPLEX, SUBUNITS, AND SUBCELLULAR LOCATION.
 RX MEDLINE=98094364; PubMed=9430634;
 RA Jungmann J., Munro S.;
 RT "Multi-protein complexes in the cis Golgi of Saccharomyces cerevisiae
 RT with alpha-1,6-mannosyltransferase activity";
 RL EMBO J. 17:423-434(1998).

[6]
RN SUBCELLULAR LOCATION.
RP MEDLINE=20558560; PubMed=11095735;
RX Todorow Z., Spang A., Carnack E., Yates J., Schekman R.;
RA "Active recycling of yeast Golgi mannosyltransferase complexes through
RT the endoplasmic reticulum";
RL Proc. Natl. Acad. Sci. U.S.A. 97:13643-13648(2000).
CC -I- FUNCTION: Involved in the organization of the secretory pathway.
CC Required to maintain a functional Golgi apparatus.
CC -I- FUNCTION: The M-Pol II complex possesses alpha-1,6-
CC mannosyltransferase activity and is probably involved in the
CC elongation of the mannan backbone of N-linked glycans on cell wall
CC and periplasmic proteins.
CC -I- SUBUNIT: Component of the M-Pol II complex composed of ANP1, MN9,
CC MN10, MN11 and HOC1.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Cis-Golgi.
CC Recycles between endoplasmic reticulum and Golgi.
CC -I- SIMILARITY: BELONGS TO THE ANP1 / MN9 / VANI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S65964; AAD13971.1; -;
DR EMBL; L22171; AAA34426.1; -;
DR EMBL; S66114; AAB28440.1; -;
DR EMBL; L22173; AAA34937.1; -;
DR EMBL; U18779; AAB65006.1; -;
DR PIR; S50508; S50508.
DR GerMOnline; 139040; -;
DR SGD; S0000762; ANP1.
DR GO; GO:0000136; C:mannosyltransferase complex; TAS.
DR GO; GO:0000009; F:alpha-1,6-mannosyltransferase activity; IDA.
DR GO; GO:0000032; P:cell wall mannosyltransferase biosynthesis; TAS.
DR GO; GO:0006487; P:N-linked glycosylation; IDA.
DR InterPro; IPR05109; Anp1.
DR Pfam; PF03452; Anp1; 1.
DR Glycoprotein; Transmembrane; Signal-anchor; Golgi stack;
KW Endoplasmic reticulum.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 27 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 28 500 LUMENAL (POTENTIAL).
FT DOMAIN 446 473 GLN-RICH.
FT CONFLICT 220 224 HHDKD -> QSGQGN (IN REF. 1).
FT CONFLICT 313 313 F -> L (IN REF. 1).
FT CONFLICT 472 500 PQGKPLDDNDKKNKKHKEVPLDPPDRN -> RRGNLLMT
FT TRTRKNILKKFH (IN REF. 1).
SQ SEQUENCE 500 AA; 58182 MW; 845B395CE548CD14 CRC64;

Query Match 42.3%; Score 44; DB 1; Length 500;
Best Local Similarity 31.6%; Pred. No. 12;
Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 VNEPHGMGLDEETAAASKD 20
Db 106 MTPPHNLIDLSFLVSDSSD 124

Search completed: March 15, 2004, 14:00:59
Job time : 6.64706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:49:12 ; Search time 29.1765 Seconds
(without alignments)
216.283 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNFFHGMLDLEETIAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriaph:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	251	12	Q9ENQ4 hepatitis a
2	104	100.0	251	12	Q9ENQ7 hepatitis a
3	104	100.0	251	12	Q9ENQ2 hepatitis a
4	104	100.0	251	12	Q9ENP2 hepatitis a
5	104	100.0	251	12	Q9ENQ6 hepatitis a
6	104	100.0	251	12	Q9ENR1 hepatitis a
7	104	100.0	251	12	Q9ENP1 hepatitis a
8	104	100.0	251	12	Q9ENQ1 hepatitis a
9	104	100.0	251	12	Q9ENQ9 hepatitis a
10	104	100.0	251	12	Q9ENNA4 hepatitis a
11	104	100.0	251	12	Q9ENP5 hepatitis a
12	104	100.0	251	12	Q9ENP7 hepatitis a
13	104	100.0	251	12	Q9ENQ5 hepatitis a
14	104	100.0	251	12	Q9ENP9 hepatitis a
15	104	100.0	251	12	Q9ENN6 hepatitis a
16	104	100.0	251	12	Q9ENN5 hepatitis a

17	104	100.0	251	12	Q9ENN7	Q9enn7 hepatitis a
18	104	100.0	1124	12	Q84780	Q84780 hepatitis a
19	104	100.0	1161	12	Q05794	Q05794 hepatitis a
20	104	100.0	2216	12	Q9WMA2	Q9wma2 hepatitis a
21	104	100.0	2218	12	Q67824	Q67824 hepatitis a
22	104	100.0	2225	12	Q9DJ32	Q9dl32 hepatitis a
23	104	100.0	2227	12	Q9WMA0	Q9wma0 hepatitis a
24	104	100.0	2227	12	Q9WMA3	Q9wma3 hepatitis a
25	104	100.0	2227	12	Q67825	Q67825 hepatitis a
26	104	100.0	2227	12	Q9WMA1	Q9wma1 hepatitis a
27	104	100.0	2227	12	Q67826	Q67826 hepatitis a
28	104	100.0	2227	12	Q8VON6	Q8von6 hepatitis a
29	104	100.0	2227	12	Q9IFH5	Q9ifh5 hepatitis a
30	104	100.0	2227	12	Q9WMA4	Q9wma4 hepatitis a
31	98	94.2	251	12	Q9ENP8	Q9enp8 hepatitis a
32	98	94.2	251	12	Q9ENN3	Q9enn3 hepatitis a
33	98	94.2	251	12	Q9ENP0	Q9enp0 hepatitis a
34	98	94.2	251	12	Q9ENR0	Q9enr0 hepatitis a
35	98	94.2	251	12	Q9ENQ3	Q9enq3 hepatitis a
36	98	94.2	251	12	Q9ENP6	Q9enp6 hepatitis a
37	98	94.2	251	12	Q9ENQ8	Q9enq8 hepatitis a
38	98	94.2	251	12	Q9ENP3	Q9enp3 hepatitis a
39	98	94.2	251	12	Q9ENN8	Q9enn8 hepatitis a
40	98	94.2	251	12	Q9ENQ0	Q9enq0 hepatitis a
41	98	94.2	2227	12	Q8QV03	Q8qv03 hepatitis a
42	97	93.3	184	12	Q87092	Q87092 simian hepa
43	97	93.3	2225	12	Q9DWR1	Q9dwr1 hepatitis a
44	95	91.3	2218	12	Q67817	Q67817 hepatitis a
45	94	90.4	251	12	Q9ENQ2	Q9enq2 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENQ4 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ4;
AC Q9ENQ4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A201;
RA Fujiwara K.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB047659; BABI2167.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFFHGMLDLEETIAANSKD 20
Db 86 KVNFFHGMLDLEETIAANSKD 105

RESULT 2

Q9ENQ7 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ7;
AC Q9ENQ7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)


```

RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
DB 86 KVNPHGMLDLEIAANSKD 105

RESULT 5
Q9ENQ6 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
[1]
RX SEQUENCE FROM N.A.
RP STRAIN=A162;
RC Fujiwara K.;
RA "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB047657; BAB12165.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
DB 86 KVNPHGMLDLEIAANSKD 105

RESULT 6
Q9ENR1 PRELIMINARY; PRT; 251 AA.
AC Q9ENR1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
[1]
RX SEQUENCE FROM N.A.
RP STRAIN=A1;
RC Fujiwara K.;
RA "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB047652; BAB12160.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;

```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		ID Q9ENQ9 PRELIMINARY; PRT; 251 AA.	
Qy	1 KVNPHGMLDLEETAAANSKD 20 	AC Q9ENQ9; 01-MAR-2001 (Tremblrel. 16, Created)	DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
Db	86 KVNPHGMLDLEETAAANSKD 105 	RC STRAIN=A159; Fujiwara K.; RA "hepatitis A virus."; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	DE Polypeptide (Fragment). OS Hepatitis A virus. OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; OC Hepatovirus. OX NCBI_TaxID=12092; RN [1]_TaxID=12092; RP SEQUENCE FROM N.A. RC STRAIN=A159; RA Fujiwara K.; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RESULT 7		Query Match 100.0%; Score 104; DB 12; Length 251; Best Local Similarity 100.0%; Pred. No. 2.4e-09; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 KVNPHGMLDLEETAAANSKD 20 	AC Q9ENQ9; 01-MAR-2001 (Tremblrel. 16, Created)	DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
Db	86 KVNPHGMLDLEETAAANSKD 105 	RC STRAIN=A159; Fujiwara K.; RA "hepatitis A virus."; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	DE Polypeptide (Fragment). OS Hepatitis A virus. OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; OC Hepatovirus. OX NCBI_TaxID=12092; RN [1]_TaxID=12092; RP SEQUENCE FROM N.A. RC STRAIN=A159; RA Fujiwara K.; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RESULT 8		Query Match 100.0%; Score 104; DB 12; Length 251; Best Local Similarity 100.0%; Pred. No. 2.4e-09; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 KVNPHGMLDLEETAAANSKD 20 	AC Q9ENQ9; 01-MAR-2001 (Tremblrel. 16, Created)	DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
Db	86 KVNPHGMLDLEETAAANSKD 105 	RC STRAIN=A159; Fujiwara K.; RA "hepatitis A virus."; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	DE Polypeptide (Fragment). OS Hepatitis A virus. OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; OC Hepatovirus. OX NCBI_TaxID=12092; RN [1]_TaxID=12092; RP SEQUENCE FROM N.A. RC STRAIN=A159; RA Fujiwara K.; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RESULT 9		Query Match 100.0%; Score 104; DB 12; Length 251; Best Local Similarity 100.0%; Pred. No. 2.4e-09; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 KVNPHGMLDLEETAAANSKD 20 	AC Q9ENQ9; 01-MAR-2001 (Tremblrel. 16, Created)	DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
Db	86 KVNPHGMLDLEETAAANSKD 105 	RC STRAIN=A159; Fujiwara K.; RA "hepatitis A virus."; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	DE Polypeptide (Fragment). OS Hepatitis A virus. OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; OC Hepatovirus. OX NCBI_TaxID=12092; RN [1]_TaxID=12092; RP SEQUENCE FROM N.A. RC STRAIN=A159; RA Fujiwara K.; RT "hepatitis A virus."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

```

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEAANSKD 20
DB 86 KVNPHGMLDLEEAANSKD 105

RESULT 12
Q9ENP7 ID Q9ENP7 PRELIMINARY; PRT; 251 AA.
AC Q9ENP7;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A306;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEAANSKD 20
DB 86 KVNPHGMLDLEEAANSKD 105

RESULT 13
Q9ENQ5 ID Q9ENQ5 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ5;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A20;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEAANSKD 20
DB 86 KVNPHGMLDLEEAANSKD 105

RESULT 14
Q9ENP9 ID Q9ENP9 PRELIMINARY; PRT; 251 AA.
AC Q9ENP9;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A303;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047664; BAB12172.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28752 MW; 7215A28AD2CA5C1A CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEAANSKD 20
DB 86 KVNPHGMLDLEEAANSKD 105

RESULT 15
Q9ENN6 ID Q9ENN6 PRELIMINARY; PRT; 251 AA.
AC Q9ENN6;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A713;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047677; BAB12185.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEAANSKD 20
DB 86 KVNPHGMLDLEEAANSKD 105

```

Db 86 KVPFHGMLDEYAANSKD 105

Search completed: March 15, 2004, 14:05:22
Job time : 30.1765 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:24:17 ; Search time 44.2353 Seconds
(without alignments)
127.748 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLEETAAAGKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	2 AAW42928	AAW42928 Immunogen
2	101	100.0	21	4 AAB69445	AAB69445 Synthetic
3	101	100.0	2227	1 AAP60066	AAP60066 Sequence
4	101	100.0	2227	2 AAR05697	AAR05697 Attenuate
5	101	100.0	2227	2 AAW34074	AAW34074 Hepatitis
6	101	100.0	2227	3 AAB18609	AAB18609 Amino aci
7	101	100.0	2227	3 AAB18607	AAB18607 Amino aci
8	101	100.0	2227	3 AAB18608	AAB18608 Amino aci
9	101	100.0	2227	5 AAE19899	AAE19899 Hepatitis
10	101	100.0	2227	5 ABG31729	ABG31729 Attenuate
11	101	100.0	2227	5 ABG31727	ABG31727 Wild-type
12	101	100.0	2227	5 ABG31728	ABG31728 Hepatitis
13	101	100.0	2227	6 ABU08640	ABU08640 Attenuat
14	101	100.0	2227	6 ABU08641	ABU08641 Attenuat
15	101	100.0	2227	6 ABU08639	ABU08639 Wild type
16	101	100.0	2227	7 ABW00350	ABW00350 Hepatitis
17	53	52.5	20	2 AAW42927	AAW42927 Immunogen
18	53	52.5	21	4 AAB69444	AAB69444 Synthetic
19	46	45.5	424	4 ABB71556	ABB71556 Drosophil
20	45.5	45.0	930	7 ADE31435	ADE31435 Plant yie
21	44.5	44.1	1712	4 ABB60536	ABB60536 Drosophil
22	44	43.6	80	5 ABP02822	ABP02822 Human ORF
23	44	43.6	401	7 ADD25153	ADD25153 Fertility
24	44	42.6	152	2 AAW55492	AAW55492 H. pylori
25	43	42.6	157	2 AAW55327	AAW55327 H. pylori

26	43	42.6	350	2	AAW27183	AAW27183 (S)-3'-hy
27	43	42.6	536	3	AAG48090	AAG48090 Arabidops
28	43	42.6	1201	2	AAW90345	AAW90345 Drosophil
29	43	42.6	1201	4	ABB58421	ABB58421 Drosophil
30	42	41.6	89	4	AAU59799	AAU59799 Propionib
31	42	41.6	89	6	ABM56318	ABM56318 Propionib
32	42	41.6	177	3	AAG48507	AAG48507 Arabidops
33	42	41.6	226	3	AAG48506	AAG48506 Arabidops
34	42	41.6	265	3	AAG48505	AAG48505 Arabidops
35	42	41.6	365	2	AAW13493	AAW13493 Pentaeryt
36	42	41.6	365	2	AAW29461	AAW29461 Enterobac
37	42	41.6	402	4	ABB60260	ABB60260 Drosophil
38	42	41.6	410	6	ABU29699	ABU29699 Protein e
39	42	41.6	416	7	ADC95924	ADC95924 E. faeciu
40	42	41.6	425	4	ABB09060	ABB09060 Thermus c
41	42	41.6	439	6	ABP78252	ABP78252 N. gonorr
42	42	41.6	445	4	AAW40763	AAW40763 Human pol
43	42	41.6	445	4	AAW40764	AAW40764 Human pol
44	42	41.6	445	4	AAW40765	AAW40765 Human pol
45	42	41.6	538	3	AAW81713	AAW81713 Streptoco

ALIGNMENTS

RESULT 1

AAW42928

ID AAW42928 standard; peptide; 20 AA.

AC AAW42928;

DT 28-APR-1998 (first entry)

DE Immunogenic Hepatitis A virus peptide YK-1328.

KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
antibody.

OS Synthetic.

OS Hepatitis A virus.

FN WO9740147-A1.

PD 30-OCT-1997.

PF 18-APR-1997; 97WO-US006891.

PR 19-APR-1996; 96US-0015644P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Fields HA, Khudyakov YE;

DR WPI; 1997-535831/49.

PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
response to HAV in a mammal or to detect the presence of antibodies
against HAV in a mammal.

PS Claim 18; Page 112; 140pp; English.

CC Peptides AAW42922-30 are immunogenic peptides corresponding to
immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
substantially similar to a portion of the amino acid sequence of the P2A
protein of HAV corresponding to amino acids 792-980. The present peptide
is derived from amino acids 931-950, and has a reactivity of 12.5% with
acute sera. Compositions containing the peptides can be used to induce an
immune response to HAV in a mammal. The peptides can also be used to
detect the presence of antibodies against HAV in mammalian serum. The
peptides can also be used to make an antibody against HAV by
administering the peptide to a mammal

XX Sequence 20 AA;

SQ

Query Match 100.0%; Score 101; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEETAAANSKOPFNNSETDL 20
Db 1 DLEETAAANSKOPFNNSETDL 20

RESULT 2
AAB69445
ID AAB69445 standard; peptide; 21 AA.

AC AAB69445;
XX 20-APR-2001 (first entry)
XX Synthetic HAV P2A peptide, SEQ ID NO: 45.
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX Hepatitis A virus.
OS Synthetic.
XX WO200105824-A2.
XX 25-JAN-2001.
XX 14-JUL-2000; 2000WO-US019267.
XX 15-JUL-1999; 99US-0144412P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Fields HA, Khudyakov YE;

XX WPI; 2001-112681/12.
XX Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines.
XX Claim 13; Page 97; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
XX immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX comprise antigenic epitopes of the major structural capsid polypeptides
XX or non-structural polypeptides of HAV with one or more glutamine
XX molecules at the carboxy end of the peptide. The peptides are used to
XX detect the presence of antibodies against HAV in mammalian serum, to
XX detect the presence of HAV in a human or animal through the binding of
XX the peptide to an antibody, to detect acute phase infection by detecting
XX IgM antibodies in mammalian serum and detecting convalescence in a
XX mammal. The peptides are used to detect or quantify HAV antibodies in
XX samples in clinical or research-based assays using immunoblotting,
XX fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX tracking of radioactive or bioluminescent markers, chromatography or
XX electrophoresis. The peptides are used to induce an immune response to
XX HAV when administered to a human or animal. Glutamine at the carboxy end
XX of the peptides enhances the IgM antibody reactivity

XX Sequence 21 AA;

Query Match 100.0%; Score 101; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEETAAANSKOPFNNSETDL 20
Db 1 DLEETAAANSKOPFNNSETDL 20

RESULT 3
AAP60066

ID AAP60066 standard; protein; 2227 AA.

XX AAP60066;

XX 25-MAR-2003 (revised)
DT 26-JUN-1991 (first entry)

XX Sequence of viral L434 polypeptide encoded by the complete nucleotide
DE sequence of the HAV genome.

XX Diagnosis; vaccine; passive immunotherapy.

XX Hepatitis A virus.

XX Key Location/Qualifiers

FT Region 1. .245

FT Region /label= P1.1A

FT Region 246. .491

FT Region /label= 1B

FT Region 492. .836

FT Region /label= 1C

FT Region 837. .980

FT Region /label= P2.2A

FT Region 981. .1076

FT Region /label= 2B

FT Region 1077. .1422

FT Region /label= 2C

FT Region 1423. .1484

FT Region /label= P3.3A

FT Region 1485. .1507

FT Region /label= 3B

FT Region 1508. .1678

FT Region /label= 3C

FT Region 1679. .2227

FT Region /label= 3D

XX EF199480-A.

XX 29-OCT-1986.

XX 03-APR-1986; 86EP-00302465.

XX 03-APR-1985; 85US-00719329.

XX (CHIR) CHIRON CORP.

XX Dina D, Porter SJ, Vannest GA, Caput D;

XX WPI; 1986-286213/44.

XX N-PSDB; AAN60080.

XX Hepatitis A virus nucleotide sequence and polypeptide - and use in prodn.
XX of vaccines and diagnostic probes.

XX Claim 5; Fig 1; 18pp; English.

XX AAN60080 and oligonucleotide fragments are useful in detection of
XX hepatitis A virus; transformed hosts may be used for expression of
XX polypeptides and fragments useful in vaccines without risk of infection
XX by the virus or in prodn. of particles which are capable of inducing
XX immunocompetent B cells for passive immunotherapy. Pref. epitope is
XX derived from AAs 445-657 or 792-848 of the HAV polypeptide sequence
XX (AAP60066). (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEETAAANSKOPFNNSETDL 20

PS Disclosure; Fig 13A-D; 66pp; English.

XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA construct (I) comprises a genome of HAV, where the genome is a human attenuated HAV genome in which a region of the 2C gene has been replaced by a corresponding region from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The region of the 2C gene from AGM-27 contained in the construct preferably encodes amino acids 120-328 of the 2C protein, amino acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (I); (2) a cell transfected with (I) or the RNA transcript of (1); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (1) or its RNA transcript, can be used as a vaccine for preventing HAV in a mammal. (I) or the RNA transcript can also be used to stimulate the production of protective antibodies in the mammal. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 6
AAB18609
ID AAB18609 standard; protein; 2227 AA.
XX
AC AAB18609;
XX
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
XX
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-00475886.
XX
PR 18-SEP-1992; 92US-00947338.
XX
PR 17-SEP-1993; 93WO-US008610.
XX
PR 17-APR-1995; 95US-00397232.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR WPI; 2000-586464/55.
XX
DR N-PSDB; AAA75478.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.
XX
PS Disclosure; Col 93-104; 72pp; English.
XX
CC The present sequence is derived from a live attenuated hepatitis A virus (HAV) of the invention, designated HAV 4380. The sequence is produced by modifying wild type HAV strain HM-174. The HAV of the invention are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is adapted to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans

CC and other primates
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 7
AAB18607
ID AAB18607 standard; protein; 2227 AA.
XX
AC AAB18607;
XX
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-00475886.
XX
PR 18-SEP-1992; 92US-00947338.
XX
PR 17-SEP-1993; 93WO-US008610.
XX
PR 17-APR-1995; 95US-00397232.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR WPI; 2000-586464/55.
XX
DR N-PSDB; AAA75478.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.
XX
PS Disclosure; Fig 6A-K; 72pp; English.
XX
CC The present sequence is derived from a wild type hepatitis A virus (HAV) strain HM-174. The sequence is modified to produce HAV which are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is adapted to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 8
AAB18608
ID AAB18608 standard; protein; 2227 AA.
XX
AC AAB18608;

XX 15-JAN-2001 (first entry)
 DT Amino acid sequence of passage 35 Hepatitis A virus called P-35.
 DE HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW P-35 virus.
 XX Hepatitis A virus.
 OS US6113912-A.
 XX 05-SEP-2000.
 PD 07-JUN-1995; 95US-00475886.
 XX 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WC-US008610.
 PR 17-APR-1995; 95US-00397232.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 PI N-PSDB; AAA75477.
 DR Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type.
 XX Disclosure; Col 67-78; 72pp; English.
 XX The present sequence is derived from passage 35 of a wild type hepatitis
 CC A virus (HAV) strain HM-174. The resulting virus is designated P-35
 CC virus. The sequence is modified to produce HAV which are adapted to
 CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
 CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
 CC as a live vaccine for prophylaxis of hepatitis A in humans and other
 CC primates
 XX Sequence 2227 AA;
 SQ Query Match 100.0%; Score 101; DB 3; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 3.5e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DLEEIAANSKDFPNMSETDL 20
 DB 931 DLEEIAANSKDFPNMSETDL 950
 RESULT 9
 AAE19899
 ID AAE19899 standard; protein; 2227 AA.
 XX AAE19899;
 AC 18-JUN-2002 (first entry)
 DT Hepatitis A virus (HAV) protein.
 DE Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
 KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
 KW Hepatitis A virus.
 OS WO200213855-A2.
 PN 21-FEB-2002.
 XX 15-AUG-2001; 2001WO-IB001808.
 PF
 XX

PR 17-AUG-2000; 2000US-0225767P.
 PR 29-AUG-2000; 2000US-0229175P.
 PR 03-NOV-2000; 2000US-00705547.
 XX (TRIP-) TRIPEP AB.
 XX Sallberg M, Hultgren C;
 PI WPI; 2002-241837/29.
 DR N-PSDB; AAD31766.
 XX Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 XX Claim 11; Page 82-87; 120pp; English.
 XX The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to a
 CC hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is hepatitis A virus (HAV) protein
 XX Sequence 2227 AA;
 SQ Query Match 100.0%; Score 101; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 3.5e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DLEEIAANSKDFPNMSETDL 20
 DB 931 DLEEIAANSKDFPNMSETDL 950
 RESULT 10
 ABG31729
 ID ABG31729 standard; protein; 2227 AA.
 XX ABG31729;
 AC 29-AUG-2003 (revised)
 DT 29-NOV-2002 (first entry)
 XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
 DE Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
 KW HAV 4380.
 KW Hepatitis A virus; strain HM-175.
 OS US6423318-B1.
 PN 23-JUL-2002.
 PD 31-AUG-2000; 2000US-00653499.
 PF 17-SEP-1993; 93WC-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 PI WPI; 2002-680946/73.
 DR N-PSDB; ABS52789.
 DR New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 XX in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 PT

XX PS Disclosure; Col 93-104; 71pp; English.

XX CC The invention relates to a polynucleotide which encodes a hepatitis A

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

CC line). The polynucleotide is useful for preparing a vaccine against

CC hepatitis A virus infection. This sequence represents an attenuated

CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to

CC standardise OS field)

XX CC

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 5; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20

DB 931 DLEEIAANSKDFPNMSETDL 950

RESULT 11

ABG31727

ID ABG31727 standard; protein; 2227 AA.

XX AC ABG31727;

XX DT 29-AUG-2003 (revised)

DT 29-NOV-2002 (first entry)

XX DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.

XX KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.

XX OS Hepatitis A virus; strain HM-175.

XX PN US6423318-B1.

XX PD 23-JUL-2002.

XX PF 31-AUG-2000; 2000US-00653499.

XX PR 17-SEP-1993; 93WO-US008610.

PR 17-APR-1995; 95US-00397232.

PR 07-JUN-1995; 95US-00475886.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX WPI; 2002-680946/73.

DR N-PSDB; ABS52787.

XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.

PS Disclosure; Fig 6; 71pp; English.

XX The invention relates to a polynucleotide which encodes a hepatitis A

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

CC line). The polynucleotide is useful for preparing a vaccine against

CC hepatitis A virus infection. This sequence represents a hepatitis A virus

CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS

CC field)

XX CC

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 5; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20

DB 931 DLEEIAANSKDFPNMSETDL 950

RESULT 12

ABG31728

ID ABG31728 standard; protein; 2227 AA.

XX AC ABG31728;

XX DT 29-NOV-2002 (first entry)

XX DE Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.

XX KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;

KW virucide; mutant; PHAV/7; mutein.

XX OS Hepatitis A virus; strain HM-175.

OS Synthetic.

XX PH Key Location/Qualifiers

FT Misc-difference 764 /note= "Wild-type Glu substituted by Val"

FT Misc-difference 821 /note= "Wild-type Asn substituted by Ser"

FT Misc-difference 963 /note= "Wild-type Lys substituted by Arg"

FT Misc-difference 1052 /note= "Wild-type Ala substituted by Val"

FT Misc-difference 1062 /note= "Wild-type Gly substituted by Ala"

FT Misc-difference 1118 /note= "Wild-type Lys substituted by Met"

FT Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"

FT Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"

FT Misc-difference 1277 /note= "Wild-type Val substituted by Ile"

FT Misc-difference 1500 /note= "Wild-type His substituted by Tyr"

FT Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"

FT Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"

XX US6423318-B1.

XX PD 23-JUL-2002.

XX PF 31-AUG-2000; 2000US-00653499.

XX PR 17-SEP-1993; 93WO-US008610.

PR 17-APR-1995; 95US-00397232.

PR 07-JUN-1995; 95US-00475886.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX WPI; 2002-680946/73.

DR N-PSDB; ABS52788.

XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.

PS Example 3; Col 67-78; 71pp; English.

XX The invention relates to a polynucleotide which encodes a hepatitis A

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

CC line). The polynucleotide is useful for preparing a vaccine against

CC hepatitis A virus infection. This sequence represents a hepatitis A virus

CC mutant strain HM-175/7 (pHAV/7) polypeptide
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 13
ABU08640
ID ABU08640 standard; protein; 2227 AA.
XX
AC ABU08640;
XX
DT 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
DE Attenuated (pass35) hepatitis A virus strain HM-175.
XX
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
OS Hepatitis A virus; strain HM-175.
XX
FN US2002176869-A1.
XX
PD 28-NOV-2002.
XX
PF 29-APR-2002; 2002US-00135988.
XX
PR 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
PA (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
DR WPI; 2003-352605/02.
DR N-PSDB; ABX93474.
XX
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
PS Example 3; Fig 6; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated (pass 35)
CC human hepatitis A virus strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 15
ABU08639
ID ABU08639 standard; protein; 2227 AA.
XX
AC ABU08639;
XX
DT 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX

RESULT 14
ABU08641
ID ABU08641 standard; protein; 2227 AA.
XX
AC ABU08641;
XX
DT 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
DE Attenuated hepatitis A virus (4380) strain HM-175.
XX
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
OS Hepatitis A virus; strain HM-175.
XX
PN US2002176869-A1.
XX
PD 28-NOV-2002.
XX
PF 29-APR-2002; 2002US-00135988.
XX
PR 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
PA (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
DR WPI; 2003-352605/02.
DR N-PSDB; ABX93475.
XX
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
PS Disclosure; Page 45-51; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated human
CC hepatitis A virus (4380) strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 15
ABU08639
ID ABU08639 standard; protein; 2227 AA.
XX
AC ABU08639;
XX
DT 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX

DE Wild type human hepatitis A virus strain HM-175.
XX
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MEC-5 cell; hepatitis infection.
XX
XX Hepatitis A virus; strain HM-175.
OS
XX
XX US2002176869-A1.
EN
XX
XX
PD 28-NOV-2002.
XX
XX
XX 29-APR-2002; 2002US-00135988.
XX
PR 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
XX (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
PI
XX
XX WPI; 2003-352605/02.
DR N-PSDB; ABX93473.
DR
XX
XX
PT New live hepatitis A virus, HAV4380, adapted to growth in MEC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
XX Disclosure; Fig 6; 70pp; English.
PS
XX
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MEC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of wild type human hepatitis A
CC virus strain HM-175. (Updated on 23-OCT-2003 to standardise OS field)
XX
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
|||
Db 931 DLEETAAANSKDFPNMSETDL 950

Search completed: March 15, 2004, 13:59:59
Job time : 45.2353 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:15:47 ; Search time 11.2941 Seconds
(without alignments)
91.421 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101
Sequence: 1 DLEEIAANSKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	3	US-08-475-886-2
2	101	100.0	2227	3	US-08-475-886-4
3	101	100.0	2227	3	US-08-475-886-6
4	101	100.0	2227	3	US-08-397-232-2
5	101	100.0	2227	3	US-08-397-232-4
6	101	100.0	2227	3	US-09-171-387-2
7	101	100.0	2227	4	US-09-653-499-2
8	101	100.0	2227	4	US-09-653-499-4
9	101	100.0	2227	4	US-09-653-499-6
10	101	100.0	2227	4	US-10-104-966-12
11	101	100.0	2227	4	US-10-135-988-2
12	101	100.0	2227	4	US-10-135-988-4
13	101	100.0	2227	4	US-10-135-988-6
14	48	47.5	294	4	US-09-134-000C-5161
15	44	43.6	907	3	US-08-938-830-26
16	44	43.6	907	3	US-09-020-222-26
17	43	42.6	1201	3	US-09-098-901-2
18	42	41.6	365	2	US-08-983-352-2
19	42	41.6	416	4	US-09-107-532A-5551
20	41	40.6	376	3	US-09-200-965-2
21	41	40.6	508	4	US-09-252-991A-18910
22	41	40.6	528	4	US-09-356-806-8
23	40.5	40.1	191	4	US-09-198-452A-1095
24	40	39.6	85	4	US-09-543-681A-4934
25	40	39.6	285	4	US-09-543-681A-5097
26	40	39.6	408	4	US-09-134-000C-6460
27	40	39.6	496	4	US-09-339-159B-28

Sequence 3648, Ap
Sequence 14, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 544, Appl
Sequence 4654, Ap
Sequence 6089, Ap
Sequence 4803, Ap
Sequence 6783, Ap
Sequence 10, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 5357, Ap
Sequence 4, Appl
Sequence 2, Appl
Sequence 7, Appl
Sequence 8, Appl

28 39.5 39.1 224 4 US-09-134-001C-3648
29 39 38.6 50 1 US-08-127-351-14
30 39 38.6 50 1 US-08-480-367B-14
31 39 38.6 50 1 US-08-487-221A-14
32 39 38.6 50 1 US-08-480-370-14
33 39 38.6 135 4 US-09-198-452A-544
34 39 38.6 168 4 US-09-134-001C-4664
35 39 38.6 212 4 US-09-543-681A-6069
36 39 38.6 216 4 US-09-134-000C-4803
37 39 38.6 288 4 US-09-107-532A-6783
38 39 38.6 331 4 US-09-803-286A-10
39 39 38.6 352 3 US-09-285-691-26
40 39 38.6 352 3 US-09-687-147-26
41 39 38.6 359 4 US-09-134-000C-5357
42 39 38.6 387 4 US-09-314-847A-4
43 39 38.6 387 4 US-09-570-778A-2
44 39 38.6 387 4 US-09-570-778A-7
45 39 38.6 387 4 US-09-570-778A-8

ALIGNMENTS

RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232

; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 5
US-08-397-232-4
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 6
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999

; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 3
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 4
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETAAANSKDFPNMSETDL 20
Db 931 DLEETAAANSKDFPNMSETDL 950

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETAAANSKDFPNMSETDL 20
Db 931 DLEETAAANSKDFPNMSETDL 950

RESULT 8
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETAAANSKDFPNMSETDL 20
Db 931 DLEETAAANSKDFPNMSETDL 950

RESULT 9
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETAAANSKDFPNMSETDL 20
Db 931 DLEETAAANSKDFPNMSETDL 950

RESULT 10
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND

; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 11
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 12
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988

; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 13
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 14
US-09-134-000C-5161
; Sequence 5161, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5161
; LENGTH: 294


```

; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5161

Query Match          47.5%; Score 48; DB 4; Length 294;
Best Local Similarity 43.8%; Pred. NO. 3;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY      2 LEEIAANSKQFFNMSE 17
      ::: |::|::|
Db      9 IEDLVAQKDYPVSVE 24

```

```

RESULT 15
US-08-938-830-26
; Sequence 26, Application US/08938830
; Patent No. 6040437
; GENERAL INFORMATION:
; APPLICANT: Lasky, Laurence A.
; APPLICANT: Dowbenko, Donald J.
; TITLE OF INVENTION: Tyrosine Phosphorylated Cleavage
; TITLE OF INVENTION: Furo-Associated Proteins (pSTPIPs)
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,830
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/798419
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: F1066P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/425-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-938-830-26
Query Match 43.6%; Score 44; DB 3; Length 907;
Best Local Similarity 47.1%; Pred. No. 54;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

Search completed: March 15, 2004, 13:26:05
Job time : 11.2941 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:23:07 ; Search time 23.1765 Seconds

(without alignments)
182.213 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLREIANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	10	US-09-171-432A-45
2	101	100.0	352	14	US-10-369-493-6377
3	101	100.0	980	14	US-10-369-493-17445
4	101	100.0	2227	9	US-09-929-955-12
5	101	100.0	2227	13	US-10-104-966-12
6	101	100.0	2227	13	US-10-135-988-2
7	101	100.0	2227	13	US-10-135-988-4
8	101	100.0	2227	13	US-10-135-988-6
9	53	52.5	20	10	US-09-171-432A-44
10	45.5	45.0	930	15	US-10-225-067-2
11	45.5	45.0	930	15	US-10-374-780A-2562
12	44	43.6	401	14	US-10-195-144-27
13	44	43.6	401	15	US-10-345-072-27
14	44	43.6	1029	15	US-10-369-493-2141
15	44	43.6	1046	15	US-10-369-493-1547

16	43.5	43.1	5107	15	US-10-369-493-6377
17	43	42.6	285	15	US-10-369-493-16887
18	43	42.6	1072	15	US-10-369-493-17445
19	42	41.6	501	15	US-10-453-763-4
20	42	41.6	538	10	US-09-769-744A-34
21	42	41.6	659	16	US-10-389-566-1962
22	42	41.6	730	15	US-10-094-749-2401
23	42	41.6	72	9	US-09-864-761-41831
24	41	40.6	387	15	US-10-369-493-2892
25	41	40.6	459	9	US-09-925-300-1440
26	41	40.6	493	15	US-10-369-493-6276
27	41	40.6	528	14	US-10-205-522-8
28	40.5	40.1	191	15	US-10-289-762-1095
29	40	39.6	76	15	US-10-369-493-10442
30	40	39.6	284	9	US-09-810-997-1
31	40	39.6	284	14	US-10-174-209-1
32	40	39.6	340	15	US-10-104-047-3319
33	40	39.6	347	15	US-10-369-493-11695
34	40	39.6	347	15	US-10-369-493-14287
35	40	39.6	347	15	US-10-369-493-14738
36	40	39.6	347	15	US-10-369-493-15221
37	40	39.6	381	9	US-09-815-242-11280
38	40	39.6	499	9	US-09-864-761-35385
39	40	39.6	569	14	US-10-156-761-12273
40	40	39.6	604	15	US-10-104-047-3464
41	40	39.6	715	14	US-10-032-585-7002
42	40	39.6	786	14	US-10-263-568-8
43	40	39.6	824	15	US-10-369-493-5121
44	40	39.6	824	15	US-10-369-493-5122
45	40	39.6	949	15	US-10-369-493-21803

ALIGNMENTS

RESULT 1

US-09-171-432A-45
; Sequence 45, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESS: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171.432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Sequence 6377, Ap
Sequence 16887, A
Sequence 17445, A
Sequence 4, Appl
Sequence 34, Appl
Sequence 1962, Ap
Sequence 2401, Ap
Sequence 41831, A
Sequence 2892, Ap
Sequence 1440, Ap
Sequence 6276, Ap
Sequence 8, Appl
Sequence 1095, Ap
Sequence 10442, A
Sequence 1, Appl
Sequence 319, Ap
Sequence 11695, A
Sequence 14287, A
Sequence 14738, A
Sequence 15221, A
Sequence 11280, A
Sequence 35385, A
Sequence 12273, A
Sequence 3464, Ap
Sequence 7002, Ap
Sequence 8, Appl
Sequence 5121, Ap
Sequence 5122, Ap
Sequence 21803, A

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1328
US-09-171-432A-45

Query Match 100.0%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 1 DLEEIAANSKDFPNMSETDL 20

RESULT 2
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match 100.0%; Score 101; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 303 DLEEIAANSKDFPNMSETDL 322

RESULT 3
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match 100.0%; Score 101; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 4
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 5
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 101; DB 13; Length 2227;

```

```
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 6
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 7
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 8
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 9
US-09-171-432A-44
; Sequence 44, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/171,432A
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/POCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1327
US-09-171-432A-44

Query Match          52.5%; Score 53; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKD 11
Db 10 DLEETAAANSKD 20

RESULT 10
US-10-225-067-2
; Sequence 2, Application US/10225067
; Publication No. US20040019925A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Creelman, Robert A.
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omalra
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E.
; TITLE OF INVENTION: BIOCHEMISTRY-RELATED POLYNUCLEOTIDES AND
; FILE REFERENCE: 51442002042
; CURRENT APPLICATION NUMBER: US/10/225,067
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-067-2

Query Match          45.0%; Score 45.5; DB 15; Length 930;
Best Local Similarity 42.3%; Pred. No. 89;
Matches 11; Conservative 4; Mismatches 2; Indels 9; Gaps 1;

QY 2 LEEIAANSKDF-----PNNSET 18
Db 153 LEEVSATSKDFVSRANGSPNGNES 178

RESULT 11
US-10-374-780A-2562
; Sequence 2562, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.
; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENOIT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; FILE REFERENCE: 16313-0136
; CURRENT APPLICATION NUMBER: US/10/195,144
; CURRENT FILING DATE: 2002-10-01
```

; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Raphanus sativum
US-10-195-144-27

Query Match 43.6%; Score 44; DB 14; Length 401;
Best Local Similarity 56.2%; Pred. No. 59;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMS 16
:|||||:|:|:
Db 9 ELDEIAALSKEFMSLS 24

RESULT 13
US-10-345-072-27
; Sequence 27, Application US/10345072
; Publication No. US20030237112A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.
; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENOIT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; APPLICANT: LAI, FANG MING
; APPLICANT: LEFOREST, MARTIN
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; FILE REFERENCE: 16313-0210
; CURRENT APPLICATION NUMBER: US/10/345,072
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: PCT/US02/22217
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Raphanus sativum
US-10-345-072-27

Query Match 43.6%; Score 44; DB 15; Length 401;
Best Local Similarity 56.2%; Pred. No. 59;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMS 16
:|||||:|:|:
Db 9 ELDEIAALSKEFMSLS 24

RESULT 14
US-10-369-493-2141
; Sequence 2141, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2141
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1029)
; OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-2141

Query Match 43.6%; Score 44; DB 15; Length 1029;
Best Local Similarity 47.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 LEEIAANSKDFPNMSET 18
:|||||:|:|:
Db 152 LOELAASSADIFEVGST 168

RESULT 15
US-10-369-493-1547
; Sequence 1547, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1547
; LENGTH: 1046
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1547

Query Match 43.6%; Score 44; DB 15; Length 1046;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 EEEIAANSKDFPNMSET 18
:|||||:|:|:
Db 733 QDIASDAKDFTNNPET 748

Search completed: March 15, 2004, 13:53:27
Job time : 23.1765 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:50:08 ; Search time 10 seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLEETAAANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: Pirl1:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	2227	1 GNNYHM	genome polyprotein
2	101	100.0	2227	1 GNNYHR	genome polyprotein
3	101	100.0	2227	1 GNNYK	genome polyprotein
4	101	100.0	2227	1 GNNYHB	genome polyprotein
5	98	97.0	2230	1 GNNYSA	genome polyprotein
6	50	49.5	736	2 D90574	hypothetical prote
7	49	48.5	400	2 F88931	protein R11q11.1
8	45.5	45.0	930	2 A94668	Argonaute (AGO1)-1
9	45.5	45.0	1621	2 T15264	hypothetical prote
10	44	43.6	740	1 F0LJHD	gag polyprotein - f
11	44	43.6	927	2 T38127	phosphoprotein - f
12	44	43.6	1046	2 S67786	hypothetical prote
13	43.5	43.1	754	1 BABOH	peptide-aspartate
14	43.5	43.1	1451	2 S65571	probable di-trans,
15	43.5	43.1	5107	2 T29144	pattern formation
16	43	42.6	152	2 A46610	partial CDS - Caen
17	43	42.6	152	2 A71904	hypothetical prote
18	43	42.6	231	2 H70407	hypothetical prote
19	43	42.6	285	2 B87426	rhodanese family p
20	43	42.6	454	2 T02100	hypothetical prote
21	43	42.6	598	2 H71336	probable cell divi
22	43	42.6	1072	2 A84112	alkaline amylopull
23	43	42.6	1201	2 T08603	kinesin-related pr
24	42.5	42.1	913	2 T15278	hypothetical prote
25	42.5	42.1	4717	2 T41581	hypothetical coile
26	42	41.6	289	2 T23342	hypothetical prote
27	42	41.6	365	2 C81050	cytochrome c oxida
28	42	41.6	365	2 F81826	probable cytochrom
29	42	41.6	538	2 G95015	ABC transporter, A

ALIGNMENTS

RESULT 1

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C;Accession: A25981
R;Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A;Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with d
A;Reference number: A25981; MUID:87061253; PMID:3023706
A;Accession: A25981
A;Molecule type: Genomic RNA
A;Residues: 1-2227 <COH>
A;Cross-references: EMBL:M14707; NID:G329582; PIDN:AAA45465.1; PID:G329583
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltran
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-245/Product: coat protein 1B #status predicted <VP2>
F;246-491/Product: coat protein 1C #status predicted <VP3>
F;492-791/Product: coat protein 1D #status predicted <VP1>
F;792-980/Product: coat protein 2A #status predicted <C2A>
F;981-1087/Product: core protein 2B #status predicted <C2B>
F;1088-1422/Product: core protein 2C #status predicted <C2C>
F;1423-1496/Product: protein 3A #status predicted <C3A>
F;1497-1519/Product: protein 3B #status predicted <C3B>
F;1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F;1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20

Db 931 DLEETAAANSKDFPNMSETDL 950

RESULT 2

GNNYHR

genome polyprotein - human hepatitis A virus
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pr
NA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C;Accession: A03903
R;Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nes
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A;Title: Primary structure and gene organization of human hepatitis A virus.

C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C;Accession: J03003
R;Paul, A.V.; Tada, H.; von der Helm, K.; Wisse, T.; Kiehn, R.; Wimmer, E.; Deinhardt, C.;
Virus Res. 8, 153-171, 1987
A;Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A;Reference number: J03003; MUID:88045071; PMID:2823500
A;Accession: J03003
A;Molecule type: genomic RNA
A;Residues: 1-2227 <PAU>
A;Cross-references: EMBL:M20273
C;Superfamily: hepatitis A virus genome polypeptide
C;Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hyc
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-246/Product: coat protein 1B #status predicted <VP2>
F;247-491/Product: coat protein 1C #status predicted <VP3>
F;492-836/Product: coat protein 1D #status predicted <VP1>
F;837-980/Product: coat protein 2A #status predicted <P2A>
F;981-1076/Product: core protein 2B #status predicted <P2B>
F;1077-1422/Product: core protein 2C #status predicted <P2C>
F;1423-1484/Product: core protein 3A #status predicted <P3A>
F;1485-1507/Product: core protein 3B #status predicted <P3B>
F;1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 3
GNMYK
Genome polypeptide - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core pr
NA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: A94149; A25914; A94508
R;Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A;Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison wit
A;Reference number: A94149; MUID:87175701; PMID:3031686
A;Accession: A94149
A;Status: nucleic acid sequence not shown
A;Molecule type: genomic RNA
A;Residues: 1-2227 <COH>
A;Cross-references: EMBL:M16632; NID:G329594; PIDN:AAA45471.1; PID:G329595
A;Note: submitted to GenBank, August 1987
C;Superfamily: hepatitis A virus genome polypeptide
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F;1-245/Product: coat protein 1A #status predicted <P1A>
F;246-491/Product: coat protein 1B #status predicted <P1B>
F;492-836/Product: coat protein 1C #status predicted <P1C>
F;837-980/Product: coat protein 2A #status predicted <P2A>
F;981-1076/Product: core protein 2B #status predicted <P2B>
F;1077-1422/Product: core protein 2C #status predicted <P2C>
F;1423-1484/Product: core protein 3A #status predicted <P3A>
F;1485-1507/Product: core protein 3B #status predicted <P3B>
F;1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 4
GNMYHB
Genome polypeptide - human hepatitis A virus (strain MBB)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C;Accession: J03003
R;Paul, A.V.; Tada, H.; von der Helm, K.; Wisse, T.; Kiehn, R.; Wimmer, E.; Deinhardt, C.;
Virus Res. 8, 153-171, 1987
A;Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A;Reference number: J03003; MUID:88045071; PMID:2823500
A;Accession: J03003
A;Molecule type: genomic RNA
A;Residues: 1-2227 <PAU>
A;Cross-references: EMBL:M20273
C;Superfamily: hepatitis A virus genome polypeptide
C;Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hyc
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-246/Product: coat protein 1B #status predicted <VP2>
F;247-491/Product: coat protein 1C #status predicted <VP3>
F;492-836/Product: coat protein 1D #status predicted <VP1>
F;837-980/Product: coat protein 2A #status predicted <P2A>
F;981-1108/Product: core protein 2B #status predicted <P2B>
F;1109-1439/Product: core protein 2C #status predicted <P2C>
F;1439-1496/Product: core protein 3A #status predicted <P3A>
F;1497-1519/Product: genome-linked protein VPg #status predicted <VPG>
F;1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 5
GNVYA
Genome polypeptide - simian hepatitis A virus (strain AGM-27)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
C;Species: simian hepatitis A virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C;Accession: A30470; S04885; S03965
R;Tsarev, S.A.
submitted to JIPID, April 1991
A;Reference number: A30470
A;Accession: A30470
A;Molecule type: genomic RNA
A;Residues: 1-2230 <TSA>
A;Cross-references: GB:D00924; NID:G222597; PIDN:BAA00766.1; PID:G222598
R;Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A;Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure a
A;Reference number: JQ1080; MUID:91311420; PMID:1649901
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
submitted to the EMBL Data Library, May 1989
A;Reference number: S04885
A;Accession: S04885
A;Molecule type: genomic RNA
A;Residues: 1750-2164 <BAL1>
A;Cross-references: EMBL:X15461; NID:G61971; PIDN:CAA33490.1; PID:G930268
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
FEBS Lett. 247, 425-428, 1989
A;Title: Variations in genome fragments coding for RNA polymerase in human and simian h
A;Reference number: S03965; MUID:89232168; PMID:2541023
A;Accession: S03965
A;Molecule type: genomic RNA
A;Residues: 1960-2164 <BAL2>
A;Cross-references: EMBL:X15461
C;Superfamily: hepatitis A virus genome polypeptide
C;Keywords: coat protein; core protein; polypeptide
F;1-27/Product: coat protein 1A #status predicted <C1A>

F;28-249/Product: coat protein 1B #status predicted <C1B>
 F;250-495/Product: coat protein 1C #status predicted <C1C>
 F;496-795/Product: coat protein 1D #status predicted <C1D>
 F;796-984/Product: core protein 2A #status predicted <C2A>
 F;985-1091/Product: core protein 2B #status predicted <C2B>
 F;1092-1426/Product: core protein 2C #status predicted <C2C>
 F;1427-1498/Product: protein 3A #status predicted <P3A>
 F;1499-1521/Product: protein 3B #status predicted <P3B>
 F;1522-1741/Product: protein 3C #status predicted <P3C>
 F;1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 97.0%; Score 98; DB 1; Length 2230;
 Best Local Similarity 95.0%; Pred. No. 6.4e-07;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAASKDFPNMSETDL 20

Db 935 DLEETASNSKDFPNMSETDL 954

RESULT 6

D90574

hypothetical protein MYPV 5000 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C;Species: Mycoplasma pulmonis

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C;Accession: D90574

R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, I.;

Nucleic Acids Res. 29, 2145-2153, 2001

A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis

A;Reference number: A99512; MUID:21267165; PMID:11353084

A;Accession: D90574

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-736 <KUR>

A;Cross-references: GB:AL445566; PID:g14099914; PIDN:CAC13673.1; GSPDB:GN00153

A;Experimental source: strain UAB CTIP

C;Genetics:

A;Gene: MYPV 5000

A;Genetic code: SGC3

Query Match

Best Local Similarity 49.5%; Score 50; DB 2; Length 736;

Best Local Similarity 55.0%; Pred. No. 8.2;

Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 DLEETAAASKDFPNMSETDL 20

Db 218 DLYSIYNEKDAPEISEDL 237

RESULT 7

F88931

protein R11G11.1 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C;Accession: F88931

R;Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A;Reference number: A75000; MUID:99069513; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: F88931

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-400 <STO>

A;Cross-references: GB:chr_v; PIDN:AAC69076.1; PID:g2384851; GSPDB:GN00023; CBSP:R11G11.

A;Note: contains similarity to C4-type zinc fingers

C;Genetics:

A;Gene: R11G11.1

A;Map position: 5

Query Match

Best Local Similarity 48.5%; Score 49; DB 2; Length 400;

Best Local Similarity 47.4%; Pred. No. 5.9;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 LEEIAANSKDFPNMSETDL 20

Db 164 LENTSINMKDFPHSKEDV 182

RESULT 8

A84668

Argonaute (AGO1)-like protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C;Accession: A84668

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: A84668

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-930 <STO>

A;Cross-references: GB:AE002093; NID:g3885334; PIDN:AAC77862.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g27040

A;Map position: 2

C;Superfamily: rabbit translation initiation factor eIF-2C

Query Match 45.0%; Score 45.5; DB 2; Length 930;

Best Local Similarity 42.3%; Pred. No. 56;

Matches 11; Conservative 4; Mismatches 2; Indels 9; Gaps 1;

QY 2 LEEIAANSKDF-----PNMSET 18

Db 153 LEEVSATSKDFVSRANGSPNGNES 178

RESULT 9

T15264

hypothetical protein F59E12.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T15264

R;Johnson, D.

submitted to the EMBL Data Library, May 1997

A;Description: The sequence of C. elegans cosmid F59E12.

A;Reference number: Z18318

A;Accession: T15264

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1621 <JOH>

A;Cross-references: EMBL:AF003386; NID:g2088833; PID:g2088843; PIDN:AAB54259.1; GSPDB:G

A;Experimental source: strain Bristol N2; clone F59E12

C;Genetics:

A;Gene: CBSP:F59E12.9

A;Map position: 2

A;Introns: 30/3; 55/1; 200/2; 299/2; 327/2; 369/3; 589/3; 860/1; 986/1; 1278/1; 1547/1

Query Match

Best Local Similarity 45.0%; Score 45.5; DB 2; Length 1621;

Best Local Similarity 42.9%; Pred. No. 1.1e+02;

Matches 9; Conservative 8; Mismatches 1; Indels 3; Gaps 1;

QY 2 LEEIAANSKDF-----PNMSETD 19

Db 985 IKEIVASAKDFMWDVDSDD 1005

RESULT 10

FOLJHD

gag polyprotein - squirrel monkey retrovirus SMRV-H

N;Contains: core protein p16; core protein p19; probable core protein p10; probable cor

C;Species: squirrel monkey retrovirus SMRV-H

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C/Accession: A31827
R/Oda, T.; Ikeda, S.; Watanabe, S.; Hataushika, M.; Akiyama, K.; Mitsunobu, F.
Virology 167, 468-476, 1988
A/Title: Molecular cloning, complete nucleotide sequence, and gene structure of the provirus of the AIDS-related virus gag polyprotein
A/Reference number: A31827; MUID:89073750; PMID:3201749
A/Accession: A31827
A/Molecule type: DNA
A/Residues: 1-740 <ODA>
A/Cross-references: GB:M23385; NID:G332626; PIDN:AAA66451.1; PID:9807672
C/Genetics:
A/Gene: gag
C/Superfamily: AIDS-related virus gag polyprotein
C/Keywords: core protein; polyprotein
F:1-163/Product: core protein p19 #status predicted <CP9>
F:164-318/Product: core protein p16 #status predicted <CP6>
F:319-648/Product: core protein p35 #status predicted <CP5>
F:649-740/Product: core protein p10 #status predicted <CP1>
Query Match 43.6%; Score 44; DB 1; Length 740;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 10; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
QY 1 DLEETIAA--NSKDFPNNMSET 18
||||| : : : :
DB 214 DLEERAAQYNPDWPOLNT 233
RESULT 11
T38127
phosphoprotein - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Aug-2002
A/Accession: T38127; A57087; S54119
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1997
A/Reference number: Z21772
A/Accession: T38127
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-927 <BAD>
A/Cross-references: EMBL:Z95334; PIDN:CAB08599.2; GSPDB:GN000066; SPDB:SPAC20G8.05C
A/Experimental source: strain 972h; cosmid c20G8
R:Fankhauser, C.; Raymond, A.; Cerutti, L.; Utzig, S.; Hofmann, K.; Simanis, V.
Cell 82, 435-444, 1995
A/Title: The Schizosaccharomyces pombe cdc15 gene is a key element in the reorganization of the mitotic spindle
A/Reference number: A57087; MUID:95360987; PMID:7634333
A/Accession: A57087
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 'MLTKSLQ', 28-927 <FAN>
A/Cross-references: GB:X86179
R:Fankhauser, C.; Raymond, A.; Cerutti, L.; Simanis, V.
submitted to the EMBL Data Library, April 1995
A/Description: The cdc15 gene is a key element in F-actin reorganisation at mitosis.
A/Reference number: S54119
A/Accession: S54119
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 'MLTKSLQ', 28-870, 'GKFIKSL', <FA2>
C/Genetics:
A/Gene: SPDB:SPAC20G8.05c; cdc15
A/Map position: 1
A/Introns: 27/3; 58/2; 871/1
C/Superfamily: fission yeast scd2 protein; SH3 homology
C/Keywords: mitosis; phosphoprotein
F:873-924/Domain: SH3 homology <SH3>
Query Match 43.6%; Score 44; DB 2; Length 927;
Best Local Similarity 47.1%; Pred. No. 97;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

F:371-404/Domain: tetrairicopeptide repeat homology <T2>
F:13,96,466,702/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 43.1%; Score 43.5; DB 1; Length 754;
Best Local Similarity 55.6%; Pred. No. 92;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
Qy 1 DLEIIAANSKDFP-NMSE 17
Db 82 DYEVLAKADFRYLSE 99
| | | | | : | | | | | : | | | | |
| | | | | : | | | | | : | | | | |

RESULT 14
S65571
Pattern formation protein GNOM - Arabidopsis thaliana
N:Alternate names: EMB30 protein
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 07-May-1999
C:Accession: S65571; S65572
R:Busch, M.; Mayer, U.; Juergens, G.
Mol. Gen. Genet. 250, 681-691, 1996
A:Title: Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure
A:Reference number: S65571; MUID:96204508; PMID:8628228
A:Accession: S65571
A:Molecule type: DNA
A:Residues: 1-1451 <BUS>
A:Cross-references: EMBL:U36433; NID:gl209632; PID:gl209633
A:Accession: S65572
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-110, 'I', 112-866, 'G', 868-1451 <BUW>
A:Cross-references: EMBL:U36432; NID:gl209630; PID:gl209631
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C:Genetics:
A:Gene: GNOM; EMB30
A:Introns: 246/3

Query Match 43.1%; Score 43.5; DB 2; Length 1451;
Best Local Similarity 63.2%; Pred. No. 2e+02;
Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 1 DLEIIAANS-KDFPNMSET 18
Db 1302 DLEIIAAGSQDYRNMEGT 1320
| | | | | : | | | | | : | | | | |
| | | | | : | | | | | : | | | | |
RESULT 15
T29144
Partial CDS - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T29144
R:Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid K11C4.
A:Reference number: 220577
A:Accession: T29144
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5107 <PAU>
A:Cross-references: EMBL:U64854; PIDN:AAB18318.1; GSPDB:GN00023; CESP:unc-68
A:Experimental source: strain Bristol N2; clone K11C4
C:Genetics:
A:Gene: CESP:unc-68
A:Map position: 5
A:Introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 1517/
/3; 3269/2; 3313/2; 3466/1; 3519/3; 3615/3; 3629/3; 3658/2; 3710/1; 3741/3; 3779/2; 3810/
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog
Query Match 43.1%; Score 43.5; DB 2; Length 5107;
Best Local Similarity 43.5%; Pred. No. 8.3e+02;
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

Qy 1 DLEIIAANS---KDFPNMSETDL 20
| | | | | : | | | | | : | | | | |
| | | | | : | | | | | : | | | | |
Db 3377 DLEIVANNNTWYSQVNPVYVDVL 3399
Search completed: March 15, 2004, 14:06:59
Job time : 10 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:26:12 ; Search time 5.64706 seconds
(without alignments)
184.415 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLEIAANSKDFPNWSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	101	100.0	2226	1 POLG_HPAV4	P26581 hepatitis a
2	101	100.0	2226	1 POLG_HPAV8	P26582 hepatitis a
3	101	100.0	2227	1 POLG_HPAVH	P08617 hepatitis a
4	101	100.0	2227	1 POLG_HPAVL	P06441 hepatitis a
5	101	100.0	2227	1 POLG_HPAVM	P13901 hepatitis a
6	98	97.0	2230	1 POLG_HPAVS	P14553 simian hepa
7	97	96.0	2226	1 POLG_HPAV2	P26580 hepatitis a
8	46	45.5	194	1 Y041 THEAC	P57674 thermoplas
9	44	43.6	740	1 GAG_SNRVH	P21411 squitrel mo
10	44	43.6	927	1 CC15 SCHPO	Q09822 schizosacch
11	43.5	43.1	754	1 ASPH_BOVIN	Q28056 bos taurus
12	43.5	43.1	1451	1 EM30_ARATH	Q42510 arabidopsis
13	43	42.6	231	1 UPFS_AQUAE	O67291 aquifex aeo
14	43	42.6	350	1 40MT_COPUA	Q91e15 optis japo
15	42	41.6	659	1 DNAK_CHLAB	Q8gh79 chlamydophi
16	41.5	41.1	1378	1 RPOB_CMAJE	Q46124 campylobact
17	41	40.6	346	1 RPAF_HAEIN	P45042 haemophilus
18	41	40.6	376	1 PT16_HUMAN	P35237 homo sapien
19	41	40.6	526	1 CLOS_CLOHI	P09870 clostridium
20	41	40.6	528	1 UDB4_HUMAN	P06133 homo sapien
21	41	40.6	548	1 HLYB_VIBCH	P15492 vibrio chol
22	40	39.6	283	1 RIAP_SOYEN	P39657 glycine max
23	40	39.6	331	1 PLUX_UREPA	Q9pqt7 ureaplasma
24	40	39.6	381	1 LDDO_HAEIN	P46454 haemophilus
25	40	39.6	482	1 GCSB_AQUAE	O67740 aquifex aeo
26	40	39.6	547	1 NLTP_HUMAN	P22307 homo sapien
27	40	39.6	569	1 HEM1_STRAW	Q82e77 streptomyce
28	40	39.6	824	1 NSB_CABEL	Q94392 caenorhabdi
29	40	39.6	1087	1 E4L3_HUMAN	Q9y212 homo sapien
30	40	39.6	1773	1 D1P2_DROME	Q9w089 drosophila
31	39.5	39.1	210	1 GRPE_STARP	Q8cp16 staphylococ
32	39.5	39.1	244	1 P29_MYCPN	P75370 mycoplasma
33	39.5	39.1	521	1 VL12_HPV04	Q07862 human papil

34	39.5	39.1	552	1 YMB1_MYCTU	Q50684 mycobacteri
35	39	38.6	221	1 YSCL_YERPE	Q00928 yersinia pe
36	39	38.6	223	1 YSCL_YEREN	Q01253 yersinia en
37	39	38.6	253	1 PCRB_METJA	Q58647 methanococc
38	39	38.6	320	1 Y149_MYCPN	P75583 mycoplasma
39	39	38.6	394	1 UXUA_GALTY	P33668 salmonella
40	39	38.6	396	1 LDDO_ECOLI	P33232 escherichia
41	39	38.6	396	1 YJ58_YEAST	P47049 saccharomyc
42	39	38.6	400	1 DDH1_DROVI	Q24731 drosophila
43	39	38.6	438	1 GUX1_AGABI	P49075 agaricus bi
44	39	38.6	467	1 MURD_BRUME	Q8y168 bruceella me
45	39	38.6	467	1 MURD_BRUSU	Q8fzp2 bruceella su

ALIGNMENTS

RESULT 1

ID	POLG_HPAV4	STANDARD;	PRT; 2226 AA.
AC	P26581;		
DT	01-AUG-1992 (Rel. 23, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]		
DE	P3D (EC 2.7.7.48)]		
OS	Hepatitis A virus (strain 43c)		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;		
OC	Hepatovirus.		
OX	NCBI_TaxID=12095;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91162758; PubMed=1705995;		
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,		
RA	Cromeans T., Jansen R.W.;		
RT	"Antigenic and genetic variation in cytopathic hepatitis A virus		
RT	variants arising during persistent infection: evidence for genetic		
RT	recombination."		
RL	J. Virol. 65:2056-2065 (1991).		
CC	- - CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +		
CC	{RNA} (N).		
CC	- - SUBUNIT: The virus capsid is composed of 60 icosahedral units,		
CC	each of which is composed of one copy each of proteins vp1, vp2,		
CC	vp3, and vp4.		
CC	- - PTM: Specific enzymatic cleavages in vivo yield mature proteins.		
CC	- - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL; M59809; AAA45469.1; -		
DR	MEROBS; C03.005; -		
DR	InterPro; IPR004004; Calici_pol_hel.		
DR	InterPro; IPR009003; Cys Ser trypsin.		
DR	InterPro; IPR000605; RNA_helicase.		
DR	InterPro; IPR007095; RNA_pol_DS_PS.		
DR	InterPro; IPR001205; RNA_pol_P3D.		
DR	InterPro; IPR007094; RNA_pol_PSVir.		
DR	InterPro; IPR008975; Viral_cap_coat.		
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.		
DR	Pfam; PF00910; RNA_helicase; 1.		
DR	PRINTS; PR00918; CALICIVIRUSNS.		
DR	Polyprotein; Coat protein; Core protein; Transferase;		
KW	Polyprotein; Coat protein; Core protein; Thiol protease.		
KW	RNA-directed RNA polymerase; Hydrolase; Thiol protease.		
FT	CHAIN 1 23		
FT	CHAIN 24 245		
FT	CHAIN 246 491		
FT	CHAIN 246 491		

```

FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 COAT PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETIAANSKDFPNMSETDL 20
DB 931 DLEETIAANSKDFPNMSETDL 950

RESULT 2
POLG_HPAV8 STANDARD; PRT; 2226 AA.
ID POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.W., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination".
RL J. Virol. 65:2056-2065(1991).
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -I- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -I- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59808; AAA45467.1; -
CC PDB; 1QA7; 15-MAY-00.
CC MR00PS; C03.005; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3B.
CC InterPro; IPR007094; RNA_pol_P5vir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSN.
CC PolyProtein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.

```

```

FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 2511392 MW; 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETIAANSKDFPNMSETDL 20
DB 931 DLEETIAANSKDFPNMSETDL 950

RESULT 3
POLG_HPAV8 STANDARD; PRT; 2227 AA.
ID POLG_HPAV8 STANDARD; PRT; 2227 AA.
AC P06617; P06443; Q81082;
DT 01-AUG-1998 (rel. 08, Created)
DT 01-AUG-1998 (rel. 08, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RA "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses."
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RA "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus."
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RA "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase."
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -I- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -I- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -I- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -I- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT

```

CC SHOWN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M14114; AAA45475.1; -

CC EMBL; M14707; AAA45465.1; -

CC EMBL; M14707; AAA45466.1; ALT_INIT.

CC EMBL; M16632; AAA45471.1; -

CC PIR; A03905; A03905.

CC PIR; A25981; GNNYHM.

CC PIR; A34149; GNNYMK.

CC PDB; 1HAV; 23-DEC-96.

CC MEROPS; C03.005; -

CC InterPro; IPR004004; Calici_pol_hel.

CC InterPro; IPR009003; Cys Ser trypsin.

CC InterPro; IPR006005; RNA helicase.

CC InterPro; IPR007095; RNA_pol_DS_PS.

CC InterPro; IPR001205; RNA_pol_P3D.

CC InterPro; IPR007094; RNA_pol_Psvir.

CC InterPro; IPR008975; Viral_cap_coat.

CC Pfam; PF00680; RNA_dep_RNA_pol; 1.

CC Pfam; PF00910; RNA_helicase; 1.

CC PRINTS; PR00918; CALICIVIRUSNS.

CC Polyprotein; Coat protein; Core protein; Transferase;

CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.

CC CHAIN 1 23

CC COAT PROTEIN VP4 (PIA).

CC COAT PROTEIN VP2 (PIB).

CC COAT PROTEIN VP3 (PIC).

CC COAT PROTEIN VP1 (PID).

CC CORE PROTEIN P2A.

CC CORE PROTEIN P2B.

CC CORE PROTEIN P2C.

CC PROBABLE PROTEIN P3A.

CC PROBABLE PROTEIN P3B.

CC PROBABLE PROTEIN P3C.

CC RNA-DIRECTED POLYMERASE 3D.

CC K -> R (IN ATTENUATED STRAIN).

CC E -> V (IN ATTENUATED STRAIN).

CC N -> S (IN ATTENUATED STRAIN).

CC A -> V (IN ATTENUATED STRAIN).

CC G -> A (IN ATTENUATED STRAIN).

CC K -> M (IN ATTENUATED STRAIN).

CC E -> K (IN ATTENUATED STRAIN).

CC F -> S (IN ATTENUATED STRAIN).

CC V -> I (IN ATTENUATED STRAIN).

CC H -> Y (IN ATTENUATED STRAIN).

CC D -> N (IN ATTENUATED STRAIN).

CC S -> T (IN ATTENUATED STRAIN).

CC SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.1e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20

Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 4

POLG HPVAVL STANDARD; PRT; 2227 AA.

AC P06441;

DT 01-JAN-1998 (Rel. 06, Created)

DT 01-JAN-1998 (Rel. 06, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase

DE P3D (EC 2.7.7.48)).

OS Hepatitis A virus (strain LA).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatocivirus.

OX NCBI_TaxID=12099;

RN 1]

RP SEQUENCE FROM N.A.

RX MEDLINE=85190549; PubMed=2986127;

RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,

RA Merryweather J., van Nest G., Dina D.;

RT "Primary structure and gene organization of human hepatitis A virus.";

RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).

CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +

CC (RNA) (N).

CC -I- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -I- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; K02990; AAA45472.1; -

CC PIR; A03903; GNNYHR.

CC MEROPS; C03.005; -

CC InterPro; IPR004004; Calici_pol_hel.

CC InterPro; IPR009003; Cys Ser trypsin.

CC InterPro; IPR006005; RNA helicase.

CC InterPro; IPR007095; RNA_pol_DS_PS.

CC InterPro; IPR001205; RNA_pol_P3D.

CC InterPro; IPR007094; RNA_pol_Psvir.

CC InterPro; IPR008975; Viral_cap_coat.

CC Pfam; PF00680; RNA_dep_RNA_pol; 1.

CC Pfam; PF00910; RNA_helicase; 1.

CC PRINTS; PR00918; CALICIVIRUSNS.

CC Polyprotein; Coat protein; Core protein; Transferase;

CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.

CC CHAIN 1 23

CC COAT PROTEIN VP4 (PIA).

CC COAT PROTEIN VP2 (PIB).

CC COAT PROTEIN VP3 (PIC).

CC COAT PROTEIN VP1 (PID).

CC CORE PROTEIN P2A.

CC CORE PROTEIN P2B.

CC CORE PROTEIN P2C.

CC PROBABLE PROTEIN P3A.

CC PROBABLE PROTEIN P3B.

CC PROBABLE PROTEIN P3C.

CC RNA-DIRECTED POLYMERASE P3D.

CC SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.1e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20

Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 5

POLG HPVAVL STANDARD; PRT; 2227 AA.

AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;

AC Q81090; Q81091; Q81092; Q81093;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)
 Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 Hepatitis A virus (strain MBB).
 Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
 NCBI_TaxID=12100;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=8045071; PubMed=2823500;
 Paul A.V., Tada H., der Helm K., Wessel T., Kiehn R., Wimmer E., Deinhardt F.;
 "The entire nucleotide sequence of the genome of human hepatitis A virus (isolate MBB).";
 Virus Res. 8:153-171(1987).
 -/- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).
 -/- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 -/- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 -/- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 -/- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 EMBL; M20273; AAA45474.1; -
 DR MEROPS; C03.005; -
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006005; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR RNA-directed RNA polymerase; Core protein; Core protein; Transferase;
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
 FT CHAIN 837 980 CORE PROTEIN P2A.
 FT CHAIN 981 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
 FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2227 AA; 251425 MW; EC983E02A7C86349 CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 101; DB 1; Length 2227;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DLEETAAASKDFPNMSETDL 20
 DB 931 DLEETAAASKDFPNMSETDL 950
 RESULT 6
 POLG_HPAVS STANDARD; PRT; 2230 AA.
 ID POLG_HPAVS

P14553;
 01-JAN-1990 (Rel. 13, Created)
 01-AUG-1992 (Rel. 23, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 Simian hepatitis A virus (strain AGM-27).
 Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
 NCBI_TaxID=12102;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=91311420; PubMed=1649901;
 Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R., Purcell R.H.;
 "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains.";
 J. Gen. Virol. 72:1677-1683(1991).
 [2]
 SEQUENCE OF 1750-2164 FROM N.A.
 MEDLINE=89232188; PubMed=2541023;
 Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A., Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
 "Variations in genome fragments coding for RNA polymerase in human and simian hepatitis A viruses.";
 FEBS Lett. 247:425-428(1989).
 -/- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).
 -/- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 -/- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 -/- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 EMBL; D00924; BAA00766.1; -
 DR EMBL; X15461; CAA33490.1; -
 DR PIR; A30470; GNNYSA.
 DR MEROPS; C03.005; -
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006005; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR Polyprotein; Coat protein; Core protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
 FT CHAIN 1 27 COAT PROTEIN VP4 (P1A).
 FT CHAIN 28 249 COAT PROTEIN VP2 (P1B).
 FT CHAIN 250 495 COAT PROTEIN VP3 (P1C).
 FT CHAIN 496 795 COAT PROTEIN VP1 (P1D).
 FT CHAIN 796 984 CORE PROTEIN P2A.
 FT CHAIN 985 1091 CORE PROTEIN P2B.
 FT CHAIN 1092 1426 CORE PROTEIN P2C.
 FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
 FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
 FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
 FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
 Query Match
 97.0%; Score 98; DB 1; Length 2230;

Best Local Similarity 95.0%; Pred. No. 6.2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||:|||||:|||||
DB 935 DLEEIAANSKDFPNMSETDL 954

RESULT 7
POLG_HPAV2 STANDARD; PRT; 2226 AA.

AC P2580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162756; PubMed=1705995;
RA Lenon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Croweans T., Jansen R.W.;
RT "Antigenic and Genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065 (1991).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59810; AAA45468.1; -.
CC MEROPS; C03.005; -.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006005; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
FT CHAIN 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;
SQ SEQUENCE

Query Match 96.0%; Score 97; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 8.9e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||:|||||:|||||
DB 931 DLEEIAANSKDFPNMSETDL 950

RESULT 8
Y041_THEAC
ID Y041_THEAC STANDARD; PRT; 194 AA.

AC P57674;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Ta0041.
GN TA0041.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Granel W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RT Nature 407:508-513 (2000).
RL Nature 407:508-513 (2000).
CC -!- SIMILARITY: Belongs to the UPF0129 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL445063; CAC11190.1; -.
CC HAMAP; MF_00265; -; 1.
DR InterPro; IPR002851; DUF133.
DR InterPro; IPR002716; PIN.
DR InterPro; IPR006596; PINC.
DR Pfam; PF01850; PIN; 1.
DR ProDom; PD013236; DUF133; 1.
DR SMART; SM00670; PINC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 194 AA; 21854 MW; 6F7CC0D844F9FF4F CRC64;
Query Match 45.5%; Score 46; DB 1; Length 194;
Best Local Similarity 50.0%; Pred. No. 5.7;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 DLEEIAANSKDFPNMSETD 19
|||||:|||||:|||||
DB 92 VEETAARTGDLNLSQTD 109

RESULT 9
GAG_SMRVH
ID GAG_SMRVH STANDARD; PRT; 740 AA.

AC P21411;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GAG polyprotein [Contains: Core protein P19; Core protein P16;
DE Probable core protein P35; Probable core protein P10].
GN GAG.
OS Squirrel monkey retrovirus (SMRV-H) (SMRV-HLB).

Simanis V.;
Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
STRAIN=972;
MEDLINE=21848401; PubMed=11859360;
Wood V., Gilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
Mooney P., Moule S., Murgall K., Murphy L., Niblett D., Ocell C.,
Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
Wielens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
Cabel C., Fuchs M., Frizc C., Holzer E., Moestl D., Hilbert H.,
Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
Eger P., Zimmermann W., Wedler H., Wambutt R., Furelle B.,
Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
Domiguez A., Revuelta J.L., Moreno S., Armstrong J., Forebush S.L.,
Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
Shpakovsky G.V., Ussery D., Barrell B.G., Nurse P.;
"The genome sequence of Schizosaccharomyces pombe.";
Nature 415:871-880(2002).
[4]
SEQUENCE OF 286-425 FROM N.A.
STRAIN=968 h90;
MEDLINE=20232868; PubMed=10759889;
Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
Hiraoka Y.;
"Large-scale screening of intracellular protein localization in living
fission yeast cells by the use of a GFP-fusion genomic DNA library.";
Genes Cells 5:169-190(2000).
-!- FUNCTION: After the onset of mitosis, forms a ring-like structure
which co-localizes with the medial actin ring. Appears to mediate
cytoskeletal rearrangements required for cytokinesis. Essential
for viability.
-!- DEVELOPMENTAL STAGE: Peaks in early mitosis before septation.
-!- DOMAIN: The N-terminal region is in a coiled coil structure.
-!- PTM: Phosphorylated.
-!- SIMILARITY: Contains 1 FCH domain.
-!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X86179; CAA60115.1; --
EMBL; Z95334; CAB08599.2; --
EMBL; AB027810; BAB87114.1; --
PIR; T38127; T38127.
HSP; P07751; ITUD.
GeneDB SPombe; SPAC20G8.05c; --
InterPro; IPR001060; Cdc15_Fes_CIP4.
InterPro; IPR001452; SH3.
Pfam; PF00611; FCH; 1.
Pfam; PF00018; SH3; 1.
PRINTS; PR00452; SH3DOMAIN.
ProDom; PD000066; SH3; 1.
SMART; SM00055; FCH; 1.
SMART; SM00326; SH3; 1.

Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R., Feldman R.A., Short J.M., Olson G.J., Swanson R.V.
The complete genome of the hyperthermophilic bacterium Aquifex aeolicus." ;
RL Nature 392:353-358(1998) .
CC -!- FUNCTION: Generates undecaprenyl pyrophosphate (UPP) from
CC isopentenyl pyrophosphate (IPP). UPP is the precursor of the
CC carrier lipid for peptidoglycan synthesis [By similarity].
CC -!- CATALYTIC ACTIVITY: Di-trans-poly-cis-decaprenyl diphosphate +
CC isopentenyl diphosphate = diphosphate + di-trans-poly-cis-undecaprenyl diphosphate.
CC -!- SIMILARITY: Belongs to the UPP synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; AF000730; AAC07254.1; -.
DR PIR; H70407; H70407.
DR InterPro; IPR001441; UPP synth.
DR Pfam; PF01255; UPP_synthetase; 1.
DR ProDom; PD003461; UPP_synth; 1.
DR TIGRfam; TIGR00055; upps; 1.
DR PROSITE; PS01066; UPP_SYNTHETASE; 1.
KW Transferase; Cell division; Cell wall; Peptidoglycan synthesis;
KW Complete proteome.
SQ SEQUENCE 231 AA; 27148 MW; 52B9DE0A442B80AE CRC64;

Query Match 42.6%; Score 43; DB 1; Length 231;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 LBIANSKDFPNM 15
::: :|||:
Db 112 MEELESKDKFKL 125

RESULT 14
40MT COPJA
ID -4OMT COPJA STANDARD; PRT; 350 AA.
AC Q9LEF5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 3'-hydroxy-N-methyl-(S)-coclaurine 4'-O-methyltransferase
DE (BC 2.1.1.116) (S'-adenosyl-L-methionine:3'-hydroxy-N-methylcoclaurine
DE 4'-O-methyltransferase) (4'-OMT).
OS Coptis japonica (Japanese goldthread).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC Ranunculaceae; Coptis.
OX NCBI_TaxID=3442;
RN [1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP MEDLINE=20390108; PubMed=10811648;
RX Morishige T., Teujita T., Yamada Y., Sato F.;
RT "Molecular characterization of the S'-adenosyl-L-methionine:
RT 3'-hydroxy-N-methylcoclaurine 4'O-methyltransferase involved in
RT isoquinoline alkaloid biosynthesis in Coptis japonica."; ;
RL J. Biol. Chem. 275:23398-23405(2000).
CC -!- FUNCTION: CATALYZES THE TRANSFER OF THE METHYL GROUP TO THE 4'-HYDROXYL GROUP OF 3'-HYDROXY-N-METHYLCOLAURINE TO FORM RETICULINE.
CC -!- CATALYTIC ACTIVITY: S'-adenosyl-L-methionine + 3'-hydroxy-N-methyl-(S)-coclaurine = S'-adenosyl-L-homocysteine + (S)-reticulene.
CC -!- PATHWAY: Conversion of 3'-hydroxy-N-methylcoclaurine to reticulene, an important intermediate in synthesizing isoquinoline alkaloids.

CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D29812; BAB08005.1; -
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR001077; O Met trans2.
DR Pfam; PF00891; Methyltransf_2; 1.
KW Transferase; Methyltransferase.
SQ SEQUENCE 350 AA; 38775 MW; 547835EBCDEF9182 CRC64;

Query Match 42.6%; Score 43; DB 1; Length 350;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEIIAANSKDFPNM 15
|||:|||||:
Db 219 DLPHVIANSYDLFNI 233

RESULT 15
DNAX CHLAB
ID DNAX CHLAB STANDARD; PRT; 659 AA.
AC Q8GH79;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Chaperone protein dnax (Heat shock protein 70) (Heat shock 70 kDa
DE protein) (HSP70).
GN DNAX.
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB7;
RX MEDLINE=22051209; PubMed=12056482;
RA Hechard C., Grepinet O., Rodolakis A.;
RT "Protection evaluation against Chlamydomophila abortus challenge by DNA
RT vaccination with a dnax-encoding plasmid in pregnant and non-pregnant
RT mice."
RL Vet. Res. 33:313-326(2002).
CC -!- FUNCTION: Acts as a chaperone (By similarity).
CC -!- INDUCTION: By stress conditions e.g. heat shock (By similarity).
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF384685; AAN7259.1; -
DR HANAP; MF_00332; -; 1.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
FT MOD_RES 201 201 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 659 AA; 71120 MW; C1CDD8C7E1ED377F CRC64;

Query Match 41.6%; Score 42; DB 1; Length 659;
Best Local Similarity 53.3%; Pred. No. 81;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 6 AANSKDFPNMSETDL 20
|||||:|||||:
Db 609 AANAQDGFNINTEDL 623

Search completed: March 15, 2004, 14:01:00
Job time : 6.64706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:49:12 ; Search time 29.1765 Seconds
(without alignments)
216.283 Million cell updates/sec

Title: US-09-171-432A-45
Perfect score: 101
Sequence: 1 DLEEIAANSKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	251	12 Q9ENC4	Q9enc4 hepatitis a
2	101	100.0	251	12 Q9ENN2	Q9enn2 hepatitis a
3	101	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
4	101	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
5	101	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
6	101	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
7	101	100.0	251	12 Q9ENC1	Q9enc1 hepatitis a
8	101	100.0	251	12 Q9ENC9	Q9enc9 hepatitis a
9	101	100.0	251	12 Q9ENNA	Q9enn4 hepatitis a
10	101	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
11	101	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
12	101	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a
13	101	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
14	101	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
15	101	100.0	251	12 Q9ENN5	Q9enn5 hepatitis a
16	101	100.0	251	12 Q9ENN7	Q9enn7 hepatitis a

17	101	100.0	1124	12 Q84780	Q84780 hepatitis a
18	101	100.0	1161	12 Q05794	Q05794 hepatitis a
19	101	100.0	2216	12 Q9WMA2	Q9wma2 hepatitis a
20	101	100.0	2218	12 Q67824	Q67824 hepatitis a
21	101	100.0	2225	12 Q9DL32	Q9dl32 hepatitis a
22	101	100.0	2227	12 Q9WMA0	Q9wma0 hepatitis a
23	101	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
24	101	100.0	2227	12 Q67825	Q67825 hepatitis a
25	101	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
26	101	100.0	2227	12 Q67826	Q67826 hepatitis a
27	101	100.0	2227	12 Q8V0N6	Q8v0n6 hepatitis a
28	101	100.0	2227	12 Q91FH5	Q91fh5 hepatitis a
29	101	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
30	95	94.1	251	12 Q9ENP8	Q9enp8 hepatitis a
31	95	94.1	251	12 Q9ENN9	Q9enn9 hepatitis a
32	95	94.1	251	12 Q9ENN3	Q9enn3 hepatitis a
33	95	94.1	251	12 Q9ENQ7	Q9enq7 hepatitis a
34	95	94.1	251	12 Q9ENF0	Q9enf0 hepatitis a
35	95	94.1	251	12 Q9ENP4	Q9enp4 hepatitis a
36	95	94.1	251	12 Q9ENQ3	Q9enq3 hepatitis a
37	95	94.1	251	12 Q9ENP6	Q9enp6 hepatitis a
38	95	94.1	251	12 Q9ENQ8	Q9enq8 hepatitis a
39	95	94.1	251	12 Q9ENP3	Q9enp3 hepatitis a
40	95	94.1	251	12 Q9ENN8	Q9enn8 hepatitis a
41	95	94.1	251	12 Q9ENQ0	Q9enq0 hepatitis a
42	95	94.1	2218	12 Q67817	Q67817 hepatitis a
43	95	94.1	2227	12 Q8QV03	Q8qv03 hepatitis a
44	94	93.1	184	12 Q87032	Q87032 simian hepa
45	94	93.1	2225	12 Q9dwr1	Q9dwr1 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENQ4 PRELIMINARY; PRT; 251 AA.
ID Q9ENC4;
AC Q9ENQ4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A201;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB047659; BAB12167.1; -
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||
Db 95 DLEEIAANSKDFPNMSETDL 114

RESULT 2

Q9ENN2 PRELIMINARY; PRT; 251 AA.
ID Q9ENN2;
AC Q9ENN2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

```
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A9;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047681; BAB12189.1; -.
FT NON TER 1 1
FT NON TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAASKDFPNMSETDL 20
DB 95 DLEETAAASKDFPNMSETDL 114

RESULT 3
Q9ENP2 ID Q9ENP2 PRELIMINARY; PRT; 251 AA.
AC Q9ENP2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A9;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON TER 1 1
FT NON TER 251 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAASKDFPNMSETDL 20
DB 95 DLEETAAASKDFPNMSETDL 114

RESULT 4
Q9ENQ6 ID Q9ENQ6 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A162;
RA Fujiwara K.;
```

```
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -.
FT NON TER 1 1
FT NON TER 251 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAASKDFPNMSETDL 20
DB 95 DLEETAAASKDFPNMSETDL 114

RESULT 5
Q9ENR1 ID Q9ENR1 PRELIMINARY; PRT; 251 AA.
AC Q9ENR1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A1;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1; -.
FT NON TER 1 1
FT NON TER 251 251
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAASKDFPNMSETDL 20
DB 95 DLEETAAASKDFPNMSETDL 114

RESULT 6
Q9ENP1 ID Q9ENP1 PRELIMINARY; PRT; 251 AA.
AC Q9ENP1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A503;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; -.
FT NON TER 1 1
FT NON TER 251 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
95 DLEEIAANSKDFPNMSETDL 114

Db 95 DLEEIAANSKDFPNMSETDL 114

RESULT 7

Q9ENN4 PRELIMINARY; PRT; 251 AA.

AC Q9ENN4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A77;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047679; BAB12187.1; -.
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
95 DLEEIAANSKDFPNMSETDL 114

Db 95 DLEEIAANSKDFPNMSETDL 114

RESULT 8

Q9ENN4 PRELIMINARY; PRT; 251 AA.

AC Q9ENN4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A159;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
95 DLEEIAANSKDFPNMSETDL 114

Db 95 DLEEIAANSKDFPNMSETDL 114

RESULT 9

Q9ENN4 PRELIMINARY; PRT; 251 AA.

AC Q9ENN4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
95 DLEEIAANSKDFPNMSETDL 114

Db 95 DLEEIAANSKDFPNMSETDL 114

Q9ENN4 PRELIMINARY; PRT; 251 AA.

AC Q9ENN4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A77;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047679; BAB12187.1; -.
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
95 DLEEIAANSKDFPNMSETDL 114

Db 95 DLEEIAANSKDFPNMSETDL 114

RESULT 10

Q9ENN4 PRELIMINARY; PRT; 251 AA.

AC Q9ENN4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
95 DLEEIAANSKDFPNMSETDL 114

Db 95 DLEEIAANSKDFPNMSETDL 114

RESULT 11

Q9ENN4 PRELIMINARY; PRT; 251 AA.

AC Q9ENN4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
95 DLEEIAANSKDFPNMSETDL 114

Db 95 DLEEIAANSKDFPNMSETDL 114

```

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A306;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 95 DLEEIAANSKDFPNMSETDL 114
| | | | | | | | | | | | | | | | | | | | | |

RESULT 12
Q9ENQ5 ID Q9ENQ5 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A20;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 95 DLEEIAANSKDFPNMSETDL 114
| | | | | | | | | | | | | | | | | | | | | |

RESULT 13
Q9ENP9 ID Q9ENP9 PRELIMINARY; PRT; 251 AA.
AC Q9ENP9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A303;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047664; BAB12172.1; -.
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 95 DLEEIAANSKDFPNMSETDL 114
| | | | | | | | | | | | | | | | | | | | | |

RESULT 14
Q9ENN6 ID Q9ENN6 PRELIMINARY; PRT; 251 AA.
AC Q9ENN6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A713;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047677; BAB12185.1; -.
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 95 DLEEIAANSKDFPNMSETDL 114
| | | | | | | | | | | | | | | | | | | | | |

RESULT 15
Q9ENN5 ID Q9ENN5 PRELIMINARY; PRT; 251 AA.
AC Q9ENN5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A75;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047678; BAB12186.1; -.
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28659 MW; 98E8EED00B2EDF10 CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
DB 95 DLEEIAANSKDFPNMSETDL 114
| | | | | | | | | | | | | | | | | | | | | |

```


Db. 95 DLEIIAANSKDFPNNSETDL 114

Search completed: March 15, 2004, 14:05:22
Job time : 29.1765 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:24:17 ; Search time 44, 2353 Seconds

(without alignments)
127.748 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRLGLSGVQSIKQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	20	2 AAW42929	Aaw42929 Immunogen
2	96	100.0	20	4 AAB69446	Aab69446 Synthetic
3	96	100.0	25	2 AAW42969	Aaw42969 Immunogen
4	96	100.0	25	4 AAB69448	Aab69448 Synthetic
5	96	100.0	2227	1 AAP60066	Aap60066 Sequence
6	96	100.0	2227	2 AAR05697	Aar05697 Attenuate
7	96	100.0	2227	2 AAW34074	Aaw34074 Hepatitis
8	96	100.0	2227	3 AAB18609	Aab18609 Amino aci
9	96	100.0	2227	3 AAB18607	Aab18607 Amino aci
10	96	100.0	2227	3 AAB18608	Aab18608 Amino aci
11	96	100.0	2227	5 AAE19899	Aae19899 Hepatitis
12	96	100.0	2227	5 AAG31729	Agg31729 Attenuate
13	96	100.0	2227	5 AAG31727	Agg31727 Wild-type
14	96	100.0	2227	5 AAG31728	Agg31728 Hepatitis
15	96	100.0	2227	6 ABU08640	Abu08640 Attenuat
16	96	100.0	2227	6 ABU08641	Abu08641 Attenuat
17	96	100.0	2227	6 ABU08639	Abu08639 Wild type
18	96	100.0	2227	7 AAW00350	Aaw00350 Hepatitis
19	44	45.8	788	6 ABR53137	Abf53137 Protein s
20	43	44.8	80	4 AAU41911	Aau41911 Propionib
21	43	44.8	80	6 ABM38430	Abm38430 Propionib
22	43	44.8	634	6 ABU20978	Abu20978 Protein e
23	43	44.8	738	5 ABP26213	Abp26213 Streptoco
24	43	44.8	922	4 AAG85023	Aag85023 Shrimp wh
25	43	44.8	2104	7 ADC00977	Adc00977 Enterohae

26	42.5	44.3	410	5	ABP28012	Abp28012 Streptoco
27	42	43.8	115	5	ABB55412	Abb55412 Lactococc
28	42	43.8	625	4	AAU49915	Aau49915 Propionib
29	42	43.8	625	6	ABM46434	Abm46434 Propionib
30	42	43.8	1035	4	AA860277	Aab60277 Saccharom
31	42	43.8	1035	6	ABR52909	Abf52909 Protein s
32	41.5	43.2	1987	7	ABU63358	Abu63358 Human Sin
33	41.5	43.2	2013	7	ABU63356	Abu63356 Human Sin
34	41.5	43.2	2014	4	AA667395	Aag67395 Amino aci
35	41.5	43.2	2014	5	AAE24137	Aae24137 Human kin
36	41.5	43.2	2014	6	ABP60434	Abp60434 Human ste
37	41.5	43.2	2014	7	ABU63359	Abu63359 Human Sin
38	41.5	43.2	2040	7	ABU63357	Abu63357 Human Sin
39	41.5	43.2	2041	6	ABP60435	Abp60435 Human ste
40	41	42.7	156	6	ABM70111	Abm70111 Photorhab
41	41	42.7	213	4	ABB59304	Abb59304 Drosophil
42	41	42.7	291	4	ABG06041	Abg06041 Novel hum
43	41	42.7	332	2	AA37809	Aay37809 Protein w
44	41	42.7	361	6	ABU42071	Abu42071 Protein e
45	41	42.7	366	2	AAW60855	Aaw60855 Mouse CD1

ALIGNMENTS

RESULT 1

AAW42929

ID 'AAW42929 standard; peptide; 20 AA.

XX

AC AAW42929;

XX 28-APR-1998 (first entry)

XX Immunogenic Hepatitis A virus peptide YK-1331.

XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.

XX Synthetic.

OS Hepatitis A virus.

XX

PN WC9740147-A1.

XX 30-OCT-1997.

PD

PF 18-APR-1997; 97WO-US006891.

XX 19-APR-1996; 96US-0015644P.

PR (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune

PT response to HAV in a mammal or to detect the presence of antibodies

PT against HAV in a mammal.

XX Claim 18; Page 112; 140pp; English.

XX Peptides AAW42922-30 are immunogenic peptides corresponding to
immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
substantially similar to a portion of the amino acid sequence of the P2A
protein of HAV corresponding to amino acids 792-980. The present peptide
is derived from amino acids 961-980, and has a reactivity of 27.1% with
acute sera. Compositions containing the peptides can be used to induce an
immune response to HAV in a mammal. The peptides can also be used to
detect the presence of antibodies against HAV in mammalian serum. The
peptides can also be used to make an antibody against HAV by
administering the peptide to a mammal

XX Sequence 20 AA;

Query Match 100.0%; Score 96; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 |||||
 DB 1 KINLADRMGLSGVQEIKEQ 20

RESULT 2
 AAB69446
 ID AAB69446 standard; peptide; 20 AA.
 XX
 AC AAB69446;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 46.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 XX antigen; major structural capsid polypeptide; HAV antibody detection.
 OS Hepatitis A virus.
 OS Synthetic.
 XX
 PN WO200105824-A2.
 XX
 PD 25-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-US019267.
 XX
 PR 15-JUL-1999; 99US-0144412P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE;
 XX
 DR WPI; 2001-112681/12.
 XX
 PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 XX
 PS Claim 13; Page 97; 130pp; English.
 XX

The present sequence is one of a number of synthetic peptides which are
 immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 comprise antigenic epitopes of the major structural capsid polypeptides
 or non-structural polypeptides of HAV with one or more glutamine
 molecules at the carboxy end of the peptide. The peptides are used to
 detect the presence of antibodies against HAV in mammalian serum, to
 detect the presence of HAV in a human or animal through the binding of
 the peptide to an antibody, to detect acute phase infection by detecting
 IgM antibodies in mammalian serum and detecting convalescence in a
 mammal. The peptides are used to detect or quantify HAV antibodies in
 samples in clinical or research-based assays using immunoblotting,
 fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 tracking of radioactive or bioluminescent markers, chromatography or
 electrophoresis. The peptides are used to induce an immune response to
 HAV when administered to a human or animal. Glutamine at the carboxy end
 of the peptides enhances the IgM antibody reactivity

Sequence 20 AA;
 Query Match 100.0%; Score 96; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 |||||
 DB 1 KINLADRMGLSGVQEIKEQ 20

RESULT 3
 AAW42969
 ID AAW42969 standard; peptide; 25 AA.
 XX
 AC AAW42969;
 XX
 DT 28-APR-1998 (first entry)
 XX
 DE Immunogenic Hepatitis A virus peptide YK-1757.
 XX
 KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
 XX antibody.
 OS Synthetic.
 OS Hepatitis A virus.
 XX
 PN WO9740147-A1.
 XX
 PD 30-OCT-1997.
 XX
 PF 18-APR-1997; 97WO-US006891.
 XX
 PR 19-APR-1996; 96US-0015644P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE;
 XX
 DR WPI; 1997-535831/49.
 XX
 PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.
 XX
 PS Claim 18; Page 112; 140pp; English.
 XX

The present immunogenic peptide corresponds to an immunogenic epitope of
 the Hepatitis A virus (HAV). The peptide is substantially similar to a
 portion of the amino acid sequence of the P2A protein of HAV
 corresponding to amino acids 792-980. Compositions containing the peptide
 can be used to induce an immune response to HAV in a mammal. The peptide
 can also be used to detect the presence of antibodies against HAV in
 mammalian serum. The peptide can also be used to make an antibody against
 HAV by administering the peptide to a mammal

Sequence 25 AA;
 Query Match 100.0%; Score 96; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 |||||
 DB 6 KINLADRMGLSGVQEIKEQ 25

RESULT 4
 AAB69448
 ID AAB69448 standard; peptide; 25 AA.
 XX
 AC AAB69448;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 48.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 XX antigen; major structural capsid polypeptide; HAV antibody detection.
 OS Hepatitis A virus.
 OS Synthetic.
 XX
 PN WO200105824-A2.

XX 25-JAN-2001.
 XX PD
 XX 14-JUL-2000; 2000WO-US019267.
 XX PF
 XX 15-JUL-1999; 99US-0144412P.
 XX PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PA
 XX Fields HA, Khudyakov YE;
 XX PI
 XX WPI; 2001-112681/12.
 XX DR
 XX Synthetic peptides used as antigen sources for enzyme immunoassays
 XX PT detecting anti-hepatitis A virus and as vaccines.
 XX PF
 XX Claim 13; Page 99; 130pp; English.
 XX PS
 XX The present sequence is one of a number of synthetic peptides which are
 XX CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 XX CC comprise antigenic epitopes of the major structural capsid polypeptides
 XX CC or non-structural polypeptides of HAV with one or more glutamine
 XX CC molecules at the carboxy end of the peptide. The peptides are used to
 XX CC detect the presence of antibodies against HAV in mammalian serum, to
 XX CC detect the presence of HAV in a human or animal through the binding of
 XX CC the peptide to an antibody, to detect acute phase infection by detecting
 XX CC IGM antibodies in mammalian serum and detecting convalescence in a
 XX CC mammal. The peptides are used to detect or quantify HAV antibodies in
 XX CC samples in clinical or research-based assays using immunoblotting,
 XX CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 XX CC tracking of radioactive or bioluminescent markers, chromatography or
 XX CC electrophoresis. The peptides are used to induce an immune response to
 XX CC HAV when administered to a human or animal. Glutamine at the carboxy end
 XX CC of the peptides enhances the IGM antibody reactivity
 XX CC
 XX SQ Sequence 25 AA;
 Query Match 100.0%; Score 96; DB 4; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 6 KINLADRMGLSGVQEIKEQ 25
 RESULT 5
 AAP60066
 ID AAP60066 standard; protein; 2227 AA.
 XX AC AAP60066;
 XX AC
 XX 25-MAR-2003 (revised)
 XX DT 26-JUN-1991 (first entry)
 XX DT
 XX Sequence of viral L434 polypeptide encoded by the complete nucleotide
 XX DE sequence of the HAV genome.
 XX DE
 XX Diagnosis; vaccine; passive immunotherapy.
 XX KW
 XX Hepatitis A virus.
 XX OS
 XX Key Location/Qualifiers
 XX FH 1. .245
 XX FT /label= P1.1A
 XX FT 246. .491
 XX FT /label= 1B
 XX FT 492. .836
 XX FT /label= 1C
 XX FT 837. .980
 XX FT /label= P2.2A
 XX FT 981. .1076
 XX FT /label= 2B

FT Region 1077. .1422
 FT /label= 2C
 FT Region 1423. .1484
 FT /label= P3.3A
 FT Region 1485. .1507
 FT /label= 3B
 FT Region 1508. .1678
 FT /label= 3C
 FT Region 1679. .2227
 FT /label= 3D
 XX
 XX EP199480-A.
 XX PN
 XX 29-OCT-1986.
 XX PD
 XX 03-APR-1986; 86EP-00302465.
 XX PF
 XX 03-APR-1985; 85US-00719329.
 XX PR
 XX (CHIR) CHIRON CORP.
 XX PA
 XX Dina D, Potter SJ, Vannest GA, Caput D;
 XX PI
 XX WPI; 1986-286213/44.
 XX DR
 XX N-PSDB; AAN60080.
 XX DR
 XX Hepatitis A virus nucleotide sequence and polypeptide - and use in prodn.
 XX PT of vaccines and diagnostic probes.
 XX PS Claim 5; Fig 1; 18pp; English.
 XX CC AAN60080 and oligonucleotide fragments are useful in detection of
 XX CC hepatitis A virus; transfected hosts may be used for expression of
 XX CC polypeptides and fragments useful in vaccines without risk of infection
 XX CC by the virus or in prodn. of particles which are capable of inducing
 XX CC immunocompetent B cells for passive immunotherapy. Pref. epitope is
 XX CC derived from AAS 445-657 or 792-848 of the HAV polypeptide sequence
 XX CC (AAP60066). (Updated on 25-MAR-2003 to correct PA field.)
 XX CC
 XX SQ Sequence 2227 AA;
 Query Match 100.0%; Score 96; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 961 KINLADRMGLSGVQEIKEQ 980
 RESULT 6
 AAR05697
 ID AAR05697 standard; protein; 2227 AA.
 XX AC AAR05697;
 XX AC
 XX 24-OCT-2003 (revised)
 XX DT 27-AUG-2003 (revised)
 XX DT 25-MAR-2003 (revised)
 XX DT 15-AUG-1990 (first entry)
 XX DT
 XX Attenuated hepatitis A virus.
 XX DE
 XX Hepatitis A virus; vaccine; attenuated.
 XX KW
 XX Hepatitis A virus; strain HM-175.
 XX OS
 XX Key Location/Qualifiers
 XX FH 1. .23
 XX FT /label= VP4 = 1A
 XX FT 24. .245
 XX FT /label= VP2 = 1B
 XX FT 246. .491

FT	/label= VP3 = 1C
FT	492..791
OS	/label= VP1 = 1D
XX	792..980
FT	/label= 2A
FT	981..1087
FT	/label= 2B
FT	1088..11422
FT	/label= 2C
FT	1423..11496
FT	/label= 3A
FT	1497..11519
FT	/label= 3B = VPg
FT	1520..11738
FT	/label= 3C
FT	1739..2227
FT	/label= 3D
XX	
PN	US4894228-A.
XX	
PD	16-JAN-1990.
XX	
PF	12-JUL-1988; 88US-00217824.
XX	
PR	19-SEP-1984; 84US-00652067.
PR	09-SEP-1986; 86US-00905146.
XX	(USSH) US DEPT HEALTH & HUMAN SERVICE.
PA	Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
PI	Daemer RJ, Gust ID;
XX	
XX	WPI; 1990-075557/10.
DR	N-PSDB; AAQ03512.
XX	Vaccine against hepatitis A virus infection - comprises novel attenuated
PT	hepatitis A virus strain.
PT	Claim 1; Fig 1; 18pp; English.
XX	
CC	The attenuated HAV is useful for inducing protective immunity against
CC	HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
CC	several nucleotide changes distributed throughout the genome, is
CC	attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC	suitable for use as an HAV vaccine. It is noted that not all the changes
CC	are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
CC	2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC	(Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
CC	standardise OS field)
XX	
SQ	Sequence 2227 AA;
	Query Match 100.0%; Score 96; DB 2; Length 2227;
	Best Local Similarity 100.0%; Pred. No. 2.9e-07;
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 KINLADRLGLSGVQEIKEQ 20
Dd	961 KINLADRLGLSGVQEIKEQ 980
RESULT 7	
AAM34074	
ID	AAW34074 standard; protein; 2227 AA.
AC	
XX	AAW34074;
AC	
XX	
DT	17-OCT-2003 (revised)
DT	27-APR-1998 (first entry)
XX	
DE	Hepatitis A virus HM-175 protein sequence.
XX	
XX	HAV; attenuation: 2C protein; 2C chimeric hepatitis A virus; infection:

KW	vaccine.
XX	
OS	Hepatitis A virus; HM-175.
XX	
FH	Key Location/Qualifiers
FT	Protein 1..23
FT	/label= VP4
FT	24..245
FT	/label= VP2
FT	246..491
FT	/label= VP3
FT	492..791
FT	/label= VP1
FT	792..980
FT	/label= 2A
FT	981..1087
FT	/label= 2B
FT	1088..1422
FT	/label= 2C
FT	1423..11496
FT	/label= 3A
FT	1497..11519
FT	/label= 3B
FT	1520..1738
FT	/label= 3C
FT	1739..2227
FT	/label= 3D
XX	
PN	WO9740166-A2.
XX	
PD	30-OCT-1997.
XX	
PF	18-APR-1997; 97WO-US006506.
XX	
PR	19-APR-1996; 96US-0015642P.
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PA	Raychaudhuri G, Emerson SU, Purcell RH;
XX	
XX	WPI; 1997-535850/49.
DR	N-PSDB; AAT93023.
XX	
PT	Human attenuated HAV genome containing simian HAV 2C gene - useful as
PT	vaccines against HAV infection.
XX	
PS	Disclosure; Fig 13A-D; 66pp; English.
XX	
CC	This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
CC	175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
CC	by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
CC	construct (I) comprises a genome of HAV, where the genome is a human
CC	attenuated HAV genome in which a region of the 2C gene has been replaced
CC	by a corresponding region from a 2C gene of a simian AGM-27 HAV genome
CC	(see AAT93024). The region of the 2C gene from AGM-27 contained in the
CC	construct preferably encodes amino acids 120-328 of the 2C protein, amino
CC	acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript
CC	of (I); (2) a cell transfected with (I) or the RNA transcript of (1); (3)
CC	a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host
CC	cell containing the HAV of (3). (I) or its RNA transcript, can be used as
CC	a vaccine for preventing HAV in a mammal. (I) or the RNA transcript can
CC	also be used to stimulate the production of protective antibodies in the
CC	mammal. (Updated on 17-OCT-2003 to standardise OS field)
XX	
SQ	Sequence 2227 AA;
	Query Match 100.0%; Score 96; DB 2; Length 2227;
	Best Local Similarity 100.0%; Pred. No. 2.9e-07;
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 KINLADRLGLSGVQEIKEQ 20
Dd	961 KINLADRLGLSGVQEIKEQ 980

```

RESULT 8
AAB18609
ID AAB18609 standard; protein; 2227 AA.
XX
XX AAB18609;
AC
XX
XX 15-JAN-2001 (first entry)
DT
DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
XX
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX HAV 4380.
XX Hepatitis A virus.
OS
XX
XX US6113912-A.
PN
XX 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-00475886.
XX
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
PI
XX WPI; 2000-586464/55.
XX N-PSDB; AAA75478.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type.
XX
XX Disclosure; Col 93-104; 72pp; English.
XX
XX The present sequence is derived from a live attenuated hepatitis A virus
XX (HAV) of the invention, designated HAV 4380. The sequence is produced by
XX modifying wild type HAV strain HM-174. The HAV of the invention are
XX adapted to growth in the human fibroblast-like cell line MRC-5. The HAV
XX is able to propagate in MRC-5 cells and retain appropriate attenuation.
XX It is useful as a live vaccine for prophylaxis of hepatitis A in humans
XX and other primates
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 9
AAB18607
ID AAB18607 standard; protein; 2227 AA.
XX
XX AAB18607;
AC
XX
XX 15-JAN-2001 (first entry)
DT
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX
XX Hepatitis A virus.
OS
XX
```

```

PN US6113912-A.
XX
XX 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-00475886.
XX
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
PI
XX WPI; 2000-586464/55.
XX N-PSDB; AAA75478.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type.
XX
XX Disclosure; Fig 6A-K; 72pp; English.
XX
XX The present sequence is derived from a wild type hepatitis A virus (HAV)
XX strain HM-174. The sequence is modified to produce HAV which are adapted
XX to growth in the human fibroblast-like cell line MRC-5. The HAV is able
XX to propagate in MRC-5 cells and retain appropriate attenuation. It is
XX useful as a live vaccine for prophylaxis of hepatitis A in humans and
XX other primates
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 10
AAB18608
ID AAB18608 standard; protein; 2227 AA.
XX
XX AAB18608;
AC
XX
XX 15-JAN-2001 (first entry)
DT
DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX P-35 virus.
XX
XX Hepatitis A virus.
OS
XX
XX US6113912-A.
PN
XX
XX 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-00475886.
XX
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
PI
XX WPI; 2000-586464/55.
XX N-PSDB; AAA75477.
XX
```

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type.
XX
PS Disclosure; Col 67-78; 72pp; English.
XX
CC The present sequence is derived from passage 35 of a wild type hepatitis
CC A virus (HAV) strain HM-174. The resulting virus is designated P-35
CC virus. The sequence is modified to produce HAV which are adapted to
CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
CC as a live vaccine for prophylaxis of hepatitis A in humans and other
CC primates
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 11
AAE19899
ID AAE19899 standard; protein; 2227 AA.
XX
AC AAE19899;
XX
DT 18-JUN-2002 (first entry)
XX
DE Hepatitis A virus (HAV) protein.
XX
KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
OS Hepatitis A virus.
XX
PN WO200213855-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001808.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2002-241837/29.
DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Claim 11; Page 82-87; 120pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is hepatitis A virus (HAV) protein
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 12
ABG31729
ID ABG31729 standard; protein; 2227 AA.
XX
AC ABG31729;
XX
DT 29-AUG-2003 (revised)
DT 29-NOV-2002 (first entry)
XX
XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
KW HAV 4380.
XX
OS Hepatitis A virus; strain HM-175.
XX
XX US6423318-B1.
XX
PD 23-JUL-2002.
XX
PF 31-AUG-2000; 2000US-00653499.
XX
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX WPI; 2002-680946/73.
DR N-PSDB; ABS52789.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX
XX Disclosure; Col 93-104; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents an attenuated
CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
XX Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 13
ABG31727
ID ABG31727 standard; protein; 2227 AA.
XX
AC ABG31727;
XX

29-AUG-2003 (revised)
29-NOV-2002 (first entry)
Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
Hepatitis A virus; strain HM-175.
US6423318-B1.
23-JUL-2002.
31-AUG-2000; 2000US-00653499.
17-SEP-1993; 93WO-US008610.
17-APR-1995; 95US-00397232.
07-JUN-1995; 95US-00475886.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
WPI; 2002-680946/73.
N-PSDB; ABS52787.
New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
in MRC-5 cells, useful for preparing a vaccine against HAV infection.
Disclosure; Fig 6; 7lpp; English.
The invention relates to a polynucleotide which encodes a hepatitis A
virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
line). The polynucleotide is useful for preparing a vaccine against
hepatitis A virus infection. This sequence represents a hepatitis A virus
strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
field)
Sequence 2227 AA;
Query Match 100.0%; Score 96; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980
RESULT 14
ABG31728
ID ABG31728 standard; protein; 2227 AA.
XX ABG31728;
XX AC
XX AC
XX 29-NOV-2002 (first entry)
XX Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast;
XX virucide; mutant; pHAV/7; mutain.
XX Hepatitis A virus; strain HM-175.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 764
FT /note= "Wild-type Glu substituted by Val"
FT Misc-difference 821
FT /note= "Wild-type Asn substituted by Ser"
FT Misc-difference 963
FT /label= Wild-type Lys substituted by Arg

Misc-difference 1052
/note= "Wild-type Ala substituted by Val"
FT Misc-difference 1062
/note= "Wild-type Gly substituted by Ala"
FT Misc-difference 1118
/note= "Wild-type Lys substituted by Met"
FT Misc-difference 1151
/note= "Wild-type Glu substituted by Lys"
FT Misc-difference 1163
/note= "Wild-type Phe substituted by Ser"
FT Misc-difference 1277
/note= "Wild-type Val substituted by Ile"
FT Misc-difference 1500
/note= "Wild-type His substituted by Tyr"
FT Misc-difference 1805
/note= "Wild-type Asp substituted by Asn"
FT Misc-difference 1930
/note= "Wild-type Ser substituted by Thr"
XX US6423318-B1.
XX 23-JUL-2002.
XX 31-AUG-2000; 2000US-00653499.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX 07-JUN-1995; 95US-00475886.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2002-680946/73.
XX N-PSDB; ABS52788.
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX Example 3; Col 67-78; 7lpp; English.
XX The invention relates to a polynucleotide which encodes a hepatitis A
virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
line). The polynucleotide is useful for preparing a vaccine against
hepatitis A virus infection. This sequence represents a hepatitis A virus
mutant strain HM-175/7 (pHAV/7) polypeptide
XX Sequence 2227 AA;
Query Match 100.0%; Score 96; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980
RESULT 15
ABU08640
ID ABU08640 standard; protein; 2227 AA.
XX AC
XX AC
XX AC
XX 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX Attenuated (pass35) hepatitis A virus strain HM-175.
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX

OS Hepatitis A virus; strain HM-175.

XX FN US2002176869-A1.

XX PD 28-NOV-2002.

XX PF 29-APR-2002; 2002US-00135988.

XX PR 18-SEP-1992; 92US-00947338.

XX PR 17-SEP-1993; 93WO-US008610.

XX PR 17-APR-1995; 95US-00397232.

XX PR 07-JUN-1995; 95US-00475886.

XX PR 31-AUG-2000; 2000US-00653499.

XX PA (FUNK/) FUNKHOUSER A W.

XX PA (EMER/) EMERSON S U.

XX PA (PURC/) PURCELL R H.

XX PA (DHON/) D'HONDT E.

XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX DR WPI; 2003-352605/02.

XX DR N-PSDB; ABX93474.

XX PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,

XX PT useful in vaccines for protecting primates against hepatitis infection

XX PT and disease.

XX PS Example 3; Fig 6; 70pp; English.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 6; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.9e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRLGLSGVQEIKEQ 20

Db 961 KINLADRLGLSGVQEIKEQ 980

Search completed: March 15, 2004, 14:00:00

Job time : 45.2353 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:15:47 ; Search time 11.2941 Seconds
(without alignments)
91.421 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRLGLSGVQEIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A.COMB.pdp.*
- 2: /cgn2_6/ptodata/2/iaa/5B.COMB.pdp.*
- 3: /cgn2_6/ptodata/2/iaa/6A.COMB.pdp.*
- 4: /cgn2_6/ptodata/2/iaa/6B.COMB.pdp.*
- 5: /cgn2_6/ptodata/2/iaa/PTCTUS.COMB.pdp.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	2227	3	US-08-475-886-2
2	96	100.0	2227	3	US-08-475-886-4
3	96	100.0	2227	3	US-08-475-886-6
4	96	100.0	2227	3	US-08-397-232-2
5	96	100.0	2227	3	US-08-397-232-4
6	96	100.0	2227	3	US-09-171-387-2
7	96	100.0	2227	4	US-09-653-499-2
8	96	100.0	2227	4	US-09-653-499-4
9	96	100.0	2227	4	US-09-653-499-6
10	96	100.0	2227	4	US-10-104-966-12
11	96	100.0	2227	4	US-10-135-988-2
12	96	100.0	2227	4	US-10-135-988-4
13	96	100.0	2227	4	US-10-135-988-6
14	42	43.8	1250	1	US-08-441-139-9
15	41	42.7	366	3	US-08-746-883-6
16	41	42.7	366	4	US-09-313-177-6
17	41	42.7	707	4	US-09-252-991A-32924
18	41	42.7	678	4	US-09-643-537-369
19	40.5	42.2	247	4	US-09-543-681A-6785
20	40.5	42.2	295	4	US-09-328-352-5192
21	40	41.7	82	4	US-09-540-236-1932
22	40	41.7	331	4	US-09-634-238-247
23	40	41.7	332	1	US-08-469-649-2
24	40	41.7	332	4	US-09-347-878-60
25	40	41.7	334	4	US-09-543-681A-6301
26	40	41.7	440	4	US-09-198-452A-44
27	40	41.7	638	2	US-08-846-762-95

28 40 41.7 643 3 US-09-178-252-25 Sequence 25, Appl
29 40 41.7 643 4 US-09-826-660-25 Sequence 25, Appl
30 40 41.7 653 4 US-09-661-322A-6 Sequence 6, Appl
31 40 41.7 1186 3 US-09-178-252-23 Sequence 23, Appl
32 40 41.7 1186 4 US-09-826-660-23 Sequence 23, Appl
33 40 41.7 1207 1 US-07-951-715A-7 Sequence 7, Appl
34 40 41.7 1207 2 US-08-459-595A-7 Sequence 7, Appl
35 40 41.7 1207 3 US-08-459-595A-7 Sequence 7, Appl
36 40 41.7 1207 3 US-08-459-595A-7 Sequence 7, Appl
37 40 41.7 1207 3 US-08-459-595A-7 Sequence 7, Appl
38 40 41.7 1207 3 US-09-053-549-8 Sequence 8, Appl
39 40 41.7 1207 4 US-09-547-422-7 Sequence 7, Appl
40 40 41.7 1227 1 US-08-448-170-8 Sequence 8, Appl
41 40 41.7 1227 3 US-09-053-549-2 Sequence 2, Appl
42 40 41.7 1227 3 US-08-961-803-9 Sequence 9, Appl
43 40 41.7 1227 4 US-09-661-322A-63 Sequence 63, Appl
44 40 41.7 1228 4 US-09-661-322A-38 Sequence 38, Appl
45 40 41.7 1229 1 US-09-100-709-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRLGLSGVQEIKEQ 20
|||
Db 961 KINLADRLGLSGVQEIKEQ 980
RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232

; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 5
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 6
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999

; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 3
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 4
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match      100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KINLADRMGLSGVQEIKEQ 20
DB      961 KINLADRMGLSGVQEIKEQ 980

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match      100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KINLADRMGLSGVQEIKEQ 20
DB      961 KINLADRMGLSGVQEIKEQ 980

RESULT 8
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-4

Query Match      100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KINLADRMGLSGVQEIKEQ 20
DB      961 KINLADRMGLSGVQEIKEQ 980

RESULT 9
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match      100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KINLADRMGLSGVQEIKEQ 20
DB      961 KINLADRMGLSGVQEIKEQ 980

RESULT 10
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
```

; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP 23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Hepatitis A virus sequence
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 11
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 12
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988

; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 13
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 14
US-08-441-139-9
; Sequence 9, Application US/08441139
; Patent No. 5773245
; GENERAL INFORMATION:
; APPLICANT: Witttrup, Dr. Karl D.
; APPLICANT: Robinson, Anne S.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530

```

US-08-745-883-6
Query Match          42.7%; Score 41; DB 3; Length 386;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 0; Indels
OY          7  RMLGSGVQEI 17
          |||:|:|:|
Db         107 RVLGSGLQEL 117

Search completed: March 15, 2004, 13:26:06
Job time : 12.2941 secs

```

Query Match 43.8%; Score 42; DB 1; Length 1250;
Best Local Similarity 42.1%; Pred. No. 68;
Matches 8: Conservative 4; Mismatches 7; Indels

Qy 2 INLADRLGLSGVQEIKEQ 20
Db 871 LNILDRTLGFVVINEIKRK 889

RESULT 15
 US-08-746-883-6
 ; Sequence 6, Application US/08746883
 ; Patent No. 6093693
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Julius, Michael H., Filipp, D.
 ; APPLICANT: Alizadeh-Khiavi, Kamel
 ; TITLE OF INVENTION: B Cell Activation
 ; NUMBER OF SEQUENCES: 13
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Blake, Cassels & Graydon
 ; STREET: Box 25, Commerce Court West
 ; CITY: Toronto
 ; COUNTRY: Canada
 ; ZIP: M5L 1A9
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4
 ; COMPUTER: COMPAQ, IBM PC compatible
 ; OPERATING SYSTEM: MS-DOS 5.1
 ; SOFTWARE: WORD PERFECT
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/746, 893
 ; FILING DATE: No. 6093693member 18, 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hunt, John C.

NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 47841/000008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 366 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

Sequence 354, App
Sequence 1672, Ap
Sequence 19622, A
Sequence 6, Appli
Sequence 2, Appli
Sequence 8, Appli
Sequence 4, Appli
Sequence 72, Appli
Sequence 10, Appl
Sequence 12087, A
Sequence 369, App
Sequence 369, App
Sequence 369, App
Sequence 369, App
Sequence 82, Appl
Sequence 4301, Ap
Sequence 6423, Ap
Sequence 6424, Ap
Sequence 34931, A
Sequence 75, Appl
Sequence 120, App
Sequence 132, App
Sequence 1317, Ap
Sequence 1320, Ap
Sequence 44, Appl
Sequence 10680, A
Sequence 6067, Ap
Sequence 95, Appl

16 42 43.8 1250 9 US-09-801-368-364
17 42 43.8 1250 15 US-10-369-493-1672
18 41.5 43.2 946 15 US-10-369-493-19622
19 41.5 43.2 1987 14 US-10-132-382-6
20 41.5 43.2 2013 14 US-10-132-382-2
21 41.5 43.2 2014 14 US-10-132-382-8
22 41.5 43.2 2040 14 US-10-132-382-4
23 41 42.7 366 9 US-09-870-759-72
24 41 42.7 366 10 US-09-751-708A-72
25 41 42.7 366 14 US-10-281-478-10
26 41 42.7 605 9 US-09-815-242-12087
27 41 42.7 708 9 US-09-735-705-369
28 41 42.7 708 9 US-09-850-716A-369
29 41 42.7 708 9 US-09-897-778-369
30 41 42.7 708 14 US-10-117-982-369
31 41 42.7 708 15 US-10-313-986-369
32 41 42.7 1309 9 US-09-862-027-82
33 40 41.7 140 9 US-09-738-626-4301
34 40 41.7 143 15 US-10-369-493-6423
35 40 41.7 154 9 US-09-864-761-34931
36 40 41.7 316 10 US-09-907-218-76
37 40 41.7 316 15 US-10-387-629-120
38 40 41.7 331 15 US-10-264-213-152
39 40 41.7 429 16 US-10-389-566-1317
40 40 41.7 429 16 US-10-389-566-1320
41 40 41.7 429 16 US-10-389-566-1320
42 40 41.7 440 15 US-10-289-762-44
43 40 41.7 549 9 US-09-815-242-10680
44 40 41.7 615 9 US-09-738-626-6067
45 40 41.7 638 14 US-10-216-209-95

ALIGNMENTS

RESULT 1
US-09-171-432A-46
; Sequence 46, Application US/09171432A
; Publication NO. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudiyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:23:07 ; Search time 23.1765 Seconds
(without alignments)
182.213 Million cell updates/sec

Title: US-09-171-432A-46
Perfect score: 96
Sequence: 1 KINLADMLGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues 809742

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	20	US-09-171-432A-46	Sequence 46, Appl
2	96	100.0	25	US-09-171-432A-48	Sequence 45, Appl
3	96	100.0	352	US-10-272-459-45	Sequence 45, Appl
4	96	100.0	980	US-10-272-459-41	Sequence 41, Appl
5	96	100.0	2227	US-09-929-955-12	Sequence 12, Appl
6	96	100.0	2227	US-10-104-966-12	Sequence 12, Appl
7	96	100.0	2227	US-10-135-988-2	Sequence 2, Appli
8	96	100.0	2227	US-10-135-988-4	Sequence 4, Appli
9	96	100.0	2227	US-10-135-988-6	Sequence 6, Appli
10	45	46.9	622	US-10-156-761-13093	Sequence 13093, A
11	44	45.8	751	US-10-369-493-9973	Sequence 9973, Ap
12	43	44.8	145	US-10-369-493-6422	Sequence 6422, Ap
13	43	44.8	145	US-10-369-493-6425	Sequence 6425, Ap
14	42	43.8	1035	US-10-369-493-21907	Sequence 21907, A
15	42	43.8	1035	US-10-389-566-2398	Sequence 2398, Ap

```

; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1331
US-09-171-432A-46

Query Match      100.0%; Score 96; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
   |||||
Db 1 KINLADRLGLSGVQEIKEQ 20
   |||||

RESULT 2
US-09-171-432A-48
; Sequence 48, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudvakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..25
; OTHER INFORMATION: /label= YK-1757
US-09-171-432A-48

Query Match      100.0%; Score 96; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
   |||||
Db 1 KINLADRLGLSGVQEIKEQ 20
   |||||

US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kda
US-10-272-459-45

Query Match      100.0%; Score 96; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
   |||||
Db 333 KINLADRLGLSGVQEIKEQ 352
   |||||

RESULT 4
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 115.5 kda
US-10-272-459-41

Query Match      100.0%; Score 96; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
   |||||
Db 961 KINLADRLGLSGVQEIKEQ 980
   |||||

RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
US-09-929-955-12
```



```

; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: METHODS OF USE THEREOF
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

```

```

Query Match      100.0%; Score 96; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 KINLADRLGLSGVQEIKEQ 20
Db      961 KINLADRLGLSGVQEIKEQ 980

```

```

RESULT 6
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US2002015124A1
; GENERAL INFORMATION:
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: METHODS OF USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

```

```

Query Match      100.0%; Score 96; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 KINLADRLGLSGVQEIKEQ 20
Db      961 KINLADRLGLSGVQEIKEQ 980

```

```

RESULT 7
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

```

```

; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

```

```

Query Match      100.0%; Score 96; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 KINLADRLGLSGVQEIKEQ 20
Db      961 KINLADRLGLSGVQEIKEQ 980

```

```

RESULT 8
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

```

```

Query Match      100.0%; Score 96; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 KINLADRLGLSGVQEIKEQ 20
Db      961 KINLADRLGLSGVQEIKEQ 980

```

```

RESULT 9
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988

```

; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 96; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 10
US-10-156-761-13093
; Sequence 13093, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13093
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13093

Query Match 46.9%; Score 45; DB 14; Length 622;
Best Local Similarity 52.6%; Pred. No. 38;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKE 19
Db 32 KSTLADRLQLTGVDORQ 50

RESULT 11
US-10-369-493-9973
; Sequence 9973, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9973
; LENGTH: 751
; TYPE: PRT
; ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-9973

Query Match 45.8%; Score 44; DB 15; Length 751;
Best Local Similarity 52.9%; Pred. No. 71;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 INLADRLGLSGVQEIKE 18
Db 314 ILDDQMTGLSGVETVR 330

RESULT 12
US-10-369-493-6422
; Sequence 6422, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6422
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6422

Query Match 44.8%; Score 43; DB 15; Length 145;
Best Local Similarity 45.0%; Pred. No. 15;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
Db 65 KINLDGHTLSIQGEQLKTE 84

RESULT 13
US-10-369-493-6425
; Sequence 6425, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6425
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

US-10-369-493-6425

Query Match
Best Local Similarity 44.8%; Score 43; DB 15; Length 145;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 65 KINLDGHTLSIQGEQELKTE 84

Search completed: March 15, 2004, 13:53:28
Job time : 24.1765 secs

US-10-369-493-21907

Query Match
Best Local Similarity 43.8%; Score 42; DB 15; Length 1035;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKE 19
Db 159 BFNVIDISLGLSKVWEVND 177

Search completed: March 15, 2004, 13:53:28
Job time : 24.1765 secs

US-10-389-566-2398

Query Match
Best Local Similarity 43.8%; Score 42; DB 16; Length 1035;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKE 19
Db 159 BFNVIDISLGLSKVWEVND 177

Search completed: March 15, 2004, 13:53:28
Job time : 24.1765 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:50:08 ; Search time 10 Seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRLMLGLSGVQEIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	2227	1 GNNYHM	genome polyprotein
2	96	100.0	2227	1 GNNYHR	genome polyprotein
3	96	100.0	2227	1 GNNYHK	genome polyprotein
4	96	100.0	2227	1 GNNYHB	genome polyprotein
5	96	100.0	2230	1 GNNYSA	genome polyprotein
6	47	49.0	145	2 B25199	heat shock 16K pro
7	47	49.0	172	2 S33416	heat shock protein
8	44	45.8	414	2 T06303	enoyl-CoA hydratase
9	44	45.8	646	2 S72609	GTP-binding membra
10	44	45.8	653	2 G70663	probable lepa - My
11	44	45.8	788	2 S67595	hypothetical prote
12	43	44.8	145	2 B24289	heat shock protein
13	43	44.8	158	2 S64321	hypothetical prote
14	43	44.8	180	2 AC0056	conserved hypotet
15	43	44.8	370	2 H70423	oxygen-independent
16	43	44.8	2104	2 D91286	hypothetical prote
17	43	44.8	2104	2 H86127	hypothetical prote
18	42.5	44.3	602	2 AB3542	gtp-binding protei
19	42	43.8	115	2 B86882	hypothetical prote
20	42	43.8	143	1 HKW41	heat shock protein
21	42	43.8	547	2 A40656	hypothetical prote
22	42	43.8	547	2 A21022	probable membrane
23	42	43.8	646	2 D91674	conserved hypotet
24	42	43.8	1035	2 S18512	cell division cont
25	42	43.8	1250	2 A39578	SSD1 protein - yea
26	41	42.7	141	2 E64368	hypothetical prote
27	41	42.7	155	2 C34965	hypothetical 17K p
28	41	42.7	155	2 S15576	ippi protein - Shi
29	41	42.7	231	2 AE2986	conserved hypotet

ALIGNMENTS

RESULT 1

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A25981
R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with d
A:Reference number: A25981; MUID:87061253; PMID:3023706
A:Accession: A25981

A:Molecule type: Genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:G329582; PIDN:AAA45465.1; PID:G329583

C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltran
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-245/Product: coat protein 1B #status predicted <VP2>
F;246-491/Product: coat protein 1C #status predicted <VP3>
F;492-791/Product: coat protein 1D #status predicted <VP1>
F;792-980/Product: core protein 2A #status predicted <C2A>
F;981-1087/Product: core protein 2B #status predicted <C2B>
F;1088-1422/Product: core protein 2C #status predicted <C2C>
F;1423-1496/Product: protein 3A #status predicted <C3A>
F;1497-1519/Product: protein 3B #status predicted <C3B>
F;1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F;1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.8e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLMLGLSGVQEIKEQ 20

DB 961 KINLADRLMLGLSGVQEIKEQ 980

RESULT 2

GNNYHR

genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pr
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nes
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.

C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C;Accession: JS0303
R;Paul, A.V.; Tada, H.; von der Helm, K.; Wessel, T.; Kiehn, R.; Wimmer, E.; Deinhardt, R.; Paul, A.V.; Tada, H., 153-171, 1987
A;Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolated from a patient with acute hepatitis A)
A;Reference number: JS0303; MUID:88045071; PMID:2823500
A;Accession: JS0303
A;Molecule type: genomic RNA
A;Residues: 1-2227 <PAU>
A;Cross-references: EMBL:M20273
C;Superfamily: hepatitis A virus genome polypeptidase; genome-linked protein; hyc
C;Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hyc
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-246/Product: coat protein 1B #status predicted <VP2>
F;247-491/Product: coat protein 1C #status predicted <VP3>
F;492-836/Product: coat protein 1D #status predicted <VP1>
F;837-980/Product: coat protein 2A #status predicted <P2A>
F;981-1108/Product: coat protein 2B #status predicted <P2B>
F;1109-1438/Product: coat protein 2C #status predicted <P2C>
F;1439-1496/Product: coat protein 3A #status predicted <P3A>
F;1497-1519/Product: genome-linked protein VPg #status predicted <VPG>
F;1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||||
DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 5
GNVSA
N;Genome polypeptide - simian hepatitis A virus (strain AGM-27)
C;Species: simian hepatitis A virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C;Accession: A30470; S04885; S03965
R;Tsarev, S.A.
submitted to JIPID, April 1991
A;Reference number: A30470
A;Accession: A30470
A;Molecule type: genomic RNA
A;Residues: 1-2230 <TSA>
A;Cross-references: GB:D00924; NID:G222597; PIDN:BR00766.1; PID:G222598
R;Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A;Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
A;Reference number: JQ1080; MUID:91311420; PMID:1649901
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
submitted to the EMBL Data Library, May 1989
A;Reference number: S04885
A;Accession: S04885
A;Molecule type: genomic RNA
A;Residues: 1750-2164 <BAL>
A;Cross-references: EMBL:X15461; NID:G61971; PIDN:CAA33490.1; PID:G930268
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
FEBS Lett. 247, 425-428, 1989
A;Title: Variations in genome fragments coding for RNA polymerase in human and simian
A;Reference number: S03965; MUID:89232168; PMID:2541023
A;Accession: S03965
A;Molecule type: genomic RNA
A;Residues: 1960-2164 <BAL2>
A;Cross-references: EMBL:X15461
C;Superfamily: hepatitis A virus genome polypeptidase
C;Keywords: coat protein; core protein; polypeptidase
F;1-27/Product: coat protein 1A #status predicted <C1A>

F:28-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:436-795/Product: coat protein 1D #status predicted <C1D>
F:796-994/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: core protein 3A #status predicted <P3A>
F:1499-1521/Product: core protein 3B #status predicted <P3B>
F:1522-1741/Product: core protein 3C #status predicted <P3C>
F:1742-2230/Product: core protein 3D #status predicted <P3D>

Query Match 100.0%; Score 96; DB 1; Length 2230;
Best Local Similarity 100.0%; Pred. No. 2.8e-07; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
DB 965 KINLADRMGLSGVQEIKEQ 984

RESULT 6
B25199
heat shock 16K protein 2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 13-Aug-1999
C:Accession: B25199
R:Jones, D.; Russnak, R.H.; Kay, R.J.; Candido, E.P.M.
J. Biol. Chem. 261, 12006-12015, 1986
A:Title: Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis elegans
A:Reference number: A92555; MUID:86304344; PMID:3017958
A:Accession: B25199
A:Molecule type: DNA
A:Residues: 1-145 <JON>
A:Cross-references: GB:M14334; NID:g156338; PIDN:AAA28071.1; PID:g156340
C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 145;
Best Local Similarity 55.6%; Pred. No. 2.2; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 6;

QY 1 KINLADRMGLSGVQEIKE 18
DB 65 KINLDRGLTSLTQGRQELK 82

RESULT 7
S33416
heat shock protein hsp20 - nematode (Nippostrongylus brasiliensis)
C:Species: Nippostrongylus brasiliensis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C:Accession: S33416
R:Tweddie, S.; Grigg, M.E.; Ingram, L.; Selkirk, M.E.
submitted to the EMBL Data Library, April 1993
A:Description: The expression of a small heat shock homologue is developmentally regulated
A:Reference number: S33416
A:Accession: S33416
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-172 <TWE>
A:Cross-references: EMBL:X71663; NID:g297865; PIDN:CAA50655.1; PID:g297866
C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 172;
Best Local Similarity 45.0%; Pred. No. 2.6; Indels 0; Gaps 0;
Matches 9; Conservative 5; Mismatches 6;

QY 1 KINLADRMGLSGVQEIKEQ 20
DB 80 KVLQDDRLTVEGMQEVKTE 99

RESULT 8
T06303

enoyl-CoA hydratase homolog F11C18.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999
C:Accession: T06303
R:Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15589
A:Accession: T06303
A:Molecule type: DNA
A:Residues: 1-414 <BEV>
A:Cross-references: EMBL:AL049607; GSPDB:GN00062; ATSP:F11C18.10
A:Experimental source: cultivar Columbia; BAC clone F11C18
C:Genetics:
A:Gene: ATSP:F11C18.10
A:Map position: 4
A:Introns: 42/3; 75/2; 108/1; 134/3; 162/3; 191/1; 219/3; 260/2; 280/3; 313/3; 347/3; 3
Query Match 45.8%; Score 44; DB 2; Length 414;
Best Local Similarity 52.6%; Pred. No. 32; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 7;

QY 1 KINLADRMGLSGVQEIKE 19
DB 260 KIELIDKYFGLDTVEIIE 278

RESULT 9
S72609
GTP-binding membrane protein lepA - Mycobacterium leprae
N:Alternate names: protein B1937 f3_81
C:Species: Mycobacterium leprae
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 02-Feb-2001
C:Accession: S72609
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Reference number: S72580
A:Accession: S72609
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-646 <SMI>
A:Cross-references: EMBL:U00016; NID:g466961; PIDN:AAA17177.1; PID:g466991
C:Genetics:
A:Gene: lepA
A:Start codon: GTG
C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homo.
C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
F:45-177/Domain: translation elongation factor Tu homology <ETU>
F:51-58/Region: nucleotide-binding motif A (P-loop)
F:174-177/Region: GTP-binding NXKD motif

Query Match 45.8%; Score 44; DB 2; Length 646;
Best Local Similarity 71.4%; Pred. No. 35; Indels 0; Gaps 0;
Matches 10; Conservative 1; Mismatches 3;

QY 1 KINLADRMGLSGV 14
DB 57 KSTLADRMQLQTVG 70

RESULT 10
G70683
probable lepA - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Feb-2001
C:Accession: G70683
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230

search completed: 11
Job time : 11 sec

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:26:12 ; Search time 5.64706 Seconds
(without alignments)
184.415 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRLGLSGVQEIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2226	1 POLG_HPAV2	P26580 hepatitis a
2	96	100.0	2226	1 POLG_HPAV4	P26581 hepatitis a
3	96	100.0	2226	1 POLG_HPAV8	P26582 hepatitis a
4	96	100.0	2227	1 POLG_HPAVH	P05617 hepatitis a
5	96	100.0	2227	1 POLG_HPAVL	P05441 hepatitis a
6	96	100.0	2227	1 POLG_HPAVM	P13901 hepatitis a
7	96	100.0	2230	1 POLG_HPAVS	P14553 simian hepa
8	48	50.0	622	1 LEPA_STRCO	Q9rdc9 streptomyc
9	47.5	49.5	603	1 LEPA_SYNEL	Q8gm20 synechococ
10	47	49.0	145	1 HS12_CAEEL	P05822 caenorhabdi
11	47	49.0	172	1 HS20_NIPER	Q07160 nippostrong
12	45	46.9	622	1 LEPA_STRAW	Q82b23 streptomyc
13	44	45.8	646	1 LEPA_MYCLE	P53530 mycobacteri
14	44	45.8	653	1 LEPA_MYCTU	P71739 mycobacteri
15	43	44.8	145	1 HS11_CAEEL	P34696 caenorhabdi
16	43	44.8	158	1 POP6_YEAST	P53218 saccharomyc
17	43	44.8	180	1 Y454_YERPE	Q8zipl yersinia pe
18	43	44.8	598	1 LEPA_NITEU	Q820h8 nitrosomona
19	42.5	44.3	602	1 LEPA_BRUME	Q8ydb8 brucella me
20	42.5	44.3	602	1 LEPA_BRUSU	Q8fv17 brucella su
21	42	43.8	143	1 HS16_CAEEL	P05681 caenorhabdi
22	42	43.8	151	1 SODC_HALRO	P81926 halocynthia
23	42	43.8	547	1 YJDB_SALTY	P26555 salmonella
24	42	43.8	621	1 Y708_CHLMU	Q9pjw8 chlamydia m
25	42	43.8	1035	1 CC68_YEAST	P32558 saccharomyc
26	42	43.8	1250	1 SSD1_YEAST	P24276 saccharomyc
27	41	42.7	141	1 NIKR_METJA	Q57969 methanococc
28	41	42.7	155	1 IPT1_SHIFL	P18008 shigella fl
29	41	42.7	313	1 CBR1_RHIME	P58332 rhizobium m
30	41	42.7	313	1 CBR2_RHIME	P56885 rhizobium m
31	41	42.7	366	1 CD14_MOUSE	P10810 mus musculu
32	41	42.7	418	1 ASB6_MOUSE	Q91zu1 mus musculu
33	41	42.7	518	1 TH14_SCHPO	P40386 s probable

ALIGNMENTS

RESULT 1

ID	POLG_HPAV2	STANDARD;	PRT;	2226 AA.
AC	P26580;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins			
DE	P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase			
DE	P3D (EC 2.7.7.48)]			
OS	Hepatitis A virus (strain 24a).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC	Hepatovirus.			
OX	NCBI_TaxID=12094;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91162758; PubMed=1705995;			
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,			
RA	Cromeans T., Jansen R.W.;			
RT	"Antigenic and Genetic Variation in cytopathic hepatitis A virus			
RT	variants arising during persistent infection: evidence for genetic			
RT	recombination."			
RL	J. Virol. 65:2056-2065 (1991).			
CC	-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +			
CC	{RNA}(N).			
CC	-!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,			
CC	each of which is composed of one copy each of proteins VP1, VP2,			
CC	VP3, and VP4.			
CC	-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.			
CC	-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
EMBL	M59810; AAA45468.1; -			
DR	MEROFS; C03.005; -			
DR	InterPro; IPR004004; Calici_pol_hel.			
DR	InterPro; IPR009003; Cys_Ser_trypsin.			
DR	InterPro; IPR000605; RNA_helicase.			
DR	InterPro; IPR007095; RNA_pol_DS_PS.			
DR	InterPro; IPR001205; RNA_pol_PSD.			
DR	InterPro; IPR007094; RNA_pol_Psvir.			
DR	InterPro; IPR008975; Viral_cap_coat.			
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.			
DR	Pfam; PF00910; RNA_helicase; 1.			
DR	PRINTS; PR00918; CALICIVIRUSN.			
DR	Polyprotein; Coat protein; Core protein; Transferase;			
KW	RNA-directed RNA polymerase; Hydrolase; Thiol protease.			
KW	RNA-directed RNA polymerase; Hydrolase; Thiol protease.			
FT	CHAIN 1 23			
FT	CHAIN 24 245			
FT	CHAIN 246 491			

FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 |||||
 DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 2
 POLG HP4V4
 ID POLG HP4V4 STANDARD; PRT; 2226 AA.
 AC P26581;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain 43C).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromeans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065 (1991).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC each of which is composed of 60 icosahedral units,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M59809; AAA45469.1; --
 DR MEROPS; C03.005; --
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006005; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00880; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSN.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).

FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2226 AA; 251107 MW; 40384CA80B09BF75 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 |||||
 DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 3
 POLG HP4V8
 ID POLG HP4V8 STANDARD; PRT; 2226 AA.
 AC P26582;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain 18f).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromeans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065 (1991).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC each of which is composed of 60 icosahedral units,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M59808; AAA45467.1; --
 DR PDB; 1QA7; 15-MAY-00.
 DR MEROPS; C03.005; --
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006005; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSN.

KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 795 900
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;
 Query Match 100.0%; Score 96; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KINLADRLGLSGVQEIKEQ 20
 DB 961 KINLADRLGLSGVQEIKEQ 980
 RESULT 4
 ID POLG_HPAVH STANDARD; PRT; 2227 AA.
 AC P08617; P06443; Q81082;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain HM-175).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12098;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wild type;
 RX MEDLINE=87061253; PubMed=3023706;
 RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
 RA Baroudy B.M.;
 RT "Complete nucleotide sequence of wild-type hepatitis A virus:
 RT comparison with different strains of hepatitis A virus and other
 RT picornaviruses.";
 RL J. Virol. 61:50-59(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Attenuated;
 RX MEDLINE=87175701; PubMed=3013686;
 RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M.,
 RA Purcell R.H.;
 RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
 RT comparison with wild-type virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
 RN [3]
 RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
 RX MEDLINE=85162289; PubMed=2984684;
 RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
 RA Purcell R.H., Feinstein S.M.;
 RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
 RT proteins and RNA polymerase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
 CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- CAUTION: REP.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
 CC SHOWN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M14114; AAA45475.1; -
 CC EMBL; M14707; AAA45465.1; -
 CC EMBL; M14707; AAA45466.1; ALT_INIT.
 CC EMBL; M16632; AAA45471.1; -
 CC EIR; A03905; A03905.
 CC EIR; A25981; GNNYHM.
 CC EIR; A94149; GNNYMK.
 CC PDB; 1HAV; 23-DEC-96.
 CC MEROPS; C03.005; -
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUS.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 FT VARIANT 77 77
 FT VARIANT 764 764
 FT VARIANT 821 821
 FT VARIANT 1052 1052
 FT VARIANT 1062 1062
 FT VARIANT 1118 1118
 FT VARIANT 1151 1151
 FT VARIANT 1163 1163
 FT VARIANT 1277 1277
 FT VARIANT 1500 1500
 FT VARIANT 1805 1805
 FT VARIANT 1930 1930
 SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;
 Query Match 100.0%; Score 96; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KINLADRLGLSGVQEIKEQ 20
 DB 961 KINLADRLGLSGVQEIKEQ 980
 RESULT 5
 ID POLG_HPAVH STANDARD; PRT; 2227 AA.
 AC P06441;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
Hepatitis A virus (strain LA).
Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=12099;
[1]
SEQUENCE FROM N.A. PubMed=2986127;
MEDLINE=85190549; PubMed=2986127;
Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
Merryweather J., van Nest G., Dina D.;
"Primary structure and gene organization of human hepatitis A virus";
Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985)
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
-!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; K02990; AAA45472.1; -
PIR; A03903; GNNVHR.
MEROPS; C03.005; -
InterPro; IPR004004; Calici_pol_hel.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR006005; RNA_helicase.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR001205; RNA_pol_P3D.
InterPro; IPR007094; RNA_pol_PSVir.
InterPro; IPR008975; Viral_cap_coat.
Pfam; PF00880; RNA_dep_RNA_pol_1.
Pfam; PF00910; RNA_helicase; 1.
PRINTS; PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
FT CHAIN 2227 AA; 99A735484CD2799C CRC64;
SQ SEQUENCE 2227 AA; 251898 MW; 99A735484CD2799C CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVQRIKEQ 20
Db 961 KINLADRMGLSGVQRIKEQ 980

RESULT 6
POLG HPVAM STANDARD; PRT: 2227 AA.
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;

01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
Hepatitis A virus (strain MBB).
Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=12100;
[1]
SEQUENCE FROM N.A. PubMed=2823500;
MEDLINE=88045071; PubMed=2823500;
Paul A.V., Tada H., der Helm K., Wessel T., Kiehn R., Wimmer E.,
Deinhardt F.;
"The entire nucleotide sequence of the genome of human hepatitis A
virus (isolate MBB).";
Virus Res. 8:153-171(1987).
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
-!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
-!- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M20273; AAA45474.1; -
MEROPS; C03.005; -
InterPro; IPR004004; Calici_pol_hel.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR006005; RNA_helicase.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR001205; RNA_pol_P3D.
InterPro; IPR007094; RNA_pol_PSVir.
InterPro; IPR008975; Viral_cap_coat.
Pfam; PF00680; RNA_dep_RNA_pol_1.
Pfam; PF00910; RNA_helicase; 1.
PRINTS; PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT CHAIN 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVQRIKEQ 20
Db 961 KINLADRMGLSGVQRIKEQ 980

RESULT 7

POLG_HPAVS STANDARD; PRT; 2230 AA.

AC P14553; 1990 (Rel. 13, Created)

DT 01-JAN-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].

OS Simian hepatitis A virus (strain AGM-27).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12102;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91311420; PubMed=1649901;

RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R., Purcell R.H.;

RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains.";

RL J. Gen. Virol. 72:1677-1683(1991).

RN [2]

RP SEQUENCE OF 1750-2164 FROM N.A.

RX MEDLINE=9232168; PubMed=2541023;

RA Balayan M.S., Kusov Y.Y., Andjapardize A.G., Tsarev S.A., Sverdllov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;

RT "Variations in genome fragments coding for RNA polymerase in human and simian hepatitis A viruses.";

RL FEBS Lett. 247:425-428(1989).

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; D00924; BAA00766.1; -.

EMBL; X15461; CAA33490.1; -.

PIR; A30470; GNNYSA.

MEROPS; C03.005; -.

InterPro; IPR004004; Calici_pol_hel.

InterPro; IPR009003; Cys_Ser_trypsin.

InterPro; IPR000605; RNA_helicase.

InterPro; IPR007095; RNA_pol_DS_PS.

InterPro; IPR001205; RNA_pol_P3D.

InterPro; IPR007094; RNA_pol_Psvir.

InterPro; IPR008975; Viral_cap_coat.

Pfam; PF00680; RNA_dep_RNA_pol_1.

Pfam; PF00910; RNA_helicase; 1.

PRINTS; PR00918; CALICIVIRUSN.

PolyProtein; Coat protein; Core protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.

CHAIN 1 27 COAT PROTEIN VP4 (P1A).

CHAIN 28 249 COAT PROTEIN VP2 (P1B).

CHAIN 250 495 COAT PROTEIN VP3 (P1C).

CHAIN 496 795 COAT PROTEIN VP3 (P1D).

CHAIN 796 984 CORE PROTEIN P2A.

CHAIN 985 1091 CORE PROTEIN P2B.

CHAIN 1092 1426 CORE PROTEIN P2C.

CHAIN 1427 1498 PROBABLE PROTEIN 3A.

CHAIN 1499 1521 PROBABLE PROTEIN 3B.

CHAIN 1522 1741 PROBABLE PROTEIN 3C.

CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.

SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2230;

Best Local Similarity 100.0%; Pred. No. 2.3e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

QY 1 KINLADRMGLSGVQVQIKQ 20

DB 965 KINLADRMGLSGVQVQIKQ 984

RESULT 8

LEPA_STRCO STANDARD; PRT; 622 AA.

AC Q9PDC9;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE GTP-binding protein lepa.

GN LEPA OR SC02562 OR SC077.29C.

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=A3(2) / M145;

RX MEDLINE=21996410; PubMed=12000953;

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kieser T., Larke L., Murphy J., Oliver K., O'Neil S., Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;

RT Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). (2002).

RL Nature 417:141-147(2002).

CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family. LepA subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AL939113; CAB66240.1; -.

HSSP; P13551; 2EFG.

HAMAP; MF_00071; -; 1.

InterPro; IPR000795; EF_GTPbind.

InterPro; IPR000640; EFG_C.

InterPro; IPR009022; EFG_III_V.

InterPro; IPR004161; EFTU_D2.

InterPro; IPR006257; LepA.

InterPro; IPR005225; Small_GTP.

InterPro; IPR009000; Translat_factor.

Pfam; PF00679; EFG_C; 1.

Pfam; PF00009; GTP_EFTU; 1.

Pfam; PF03144; GTP_EFTU_D2; 1.

PRINTS; PR00315; ELONGATNFCT.

TIGRfams; TIGR01393; lepa; 1.

TIGRfams; TIGR00231; small_GTP; 1.

PROSITE; PS00301; EFATOR_GTP; 1.

GTP-binding; Complete proteome.

NP_BIND 26 33 GTP (BY SIMILARITY).

NP_BIND 94 98 GTP (BY SIMILARITY).

NP_BIND 148 151 GTP (BY SIMILARITY).

SEQUENCE 622 AA; 68378 MW; 83F5C76FA2A80C7C CRC64;

Query Match 50.0%; Score 48; DB 1; Length 622;
Best Local Similarity 57.9%; Pred. No. 4.6;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKE 19
DB 32 KSTLADRLQLTGTVQVEQK 50

RESULT 9
LEPA SYNELE STANDARD; PRT; 603 AA.

AC Q8DM20;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR TL0304
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.,
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1";
RL DNA Res. 9:123-130(2002).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC LePA subfamily.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AF005369; BAC07857.1; -.
DR HAMAP; MF_00071; -; 1.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR006297; LepA.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFC.
DR TIGRFAMs; TIGR01393; lepa; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACTOR_GTP; 1.
KW GTP-binding; Complete proteome.
FT NP_BIND 15 23 GTP (BY SIMILARITY).
FT NP_BIND 82 86 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 603 AA; 67578 MW; ACBBACD1407759AB CRC64;

Query Match 49.5%; Score 47.5; DB 1; Length 603;
Best Local Similarity 52.2%; Pred. No. 5.3;
Matches 12; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 1 KINLADRLGLSGV---QBIKEQ 20
DB 22 KSTLADRLQLTGTVDPREMEQ 44

RESULT 10
HS12 CAEEL STANDARD; PRT; 145 AA.

AC P06582;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heat shock protein HSP16-2.
GN HSP16-2 OR Y46H3A.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8630434; PubMed=3017958;
RA Jones D., Rusanak R.H., Kay R.J., Candido E.P.M.;
RT "Structure, expression, and evolution of a heat shock gene locus in
Caenorhabditis elegans that is flanked by repetitive elements";
RL J. Biol. Chem. 261:12006-12015(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Stoneking T., Wohlmann P., Lennox S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; M14334; AAA28071.1; -.
DR EMBL; AC006774; AAR6015.1; -.
DR PIR; B25199; B25199.
DR WormPep; Y46H3A.3; CE22002.
DR InterPro; IPR001436; Crystallin_alpha.
DR InterPro; IPR002068; Hsp20.
DR InterPro; IPR008978; HSP20_chap.
DR Pfam; PF00011; HSP20; 1.
DR PRINTS; PR00299; ACRYSTALLIN.
DR PROSITE; PS01031; HSP20; 1.
KW Heat shock; Multigene family.
SQ SEQUENCE 145 AA; 16242 MW; 8A73449F99161889 CRC64;

Query Match 49.0%; Score 47; DB 1; Length 145;
Best Local Similarity 55.6%; Pred. No. 1.5;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKE 18
DB 65 KINLADRLGLSGVQEIKE 82

RESULT 11
HS20 NIPBR STANDARD; PRT; 172 AA.

AC Q07160;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Heat shock protein homolog (HSP20).
GN HSP20.
OS Nippostrongylus brasiliensis.

RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus."
 RL Nature 409:1007-1011(2001).
 CC -1- MISCELLANEOUS: THE OVERPRODUCTION OF THIS PROTEIN IS LETHAL TO
 CC M.LEPRAE.
 CC
 CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.
 CC Lepa subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U00016; AAA17177.1; -
 CC EMBL; AL583919; CAC30119.1; -
 CC PIR; S72609; S72609.
 CC HSP; P13551; 2EFG.
 CC Leptoma; ML0611; -
 CC HAMAP; MF 00071; -
 CC InterPro; IPR000795; EF_GTPbind.
 CC InterPro; IPR000640; EFG_C.
 CC InterPro; IPR009022; EFG_III_V.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR006297; LepA.
 CC InterPro; IPR005225; Small GTP.
 CC InterPro; IPR009000; Translat_factor.
 CC Pfam; PF00679; EFG_C; 1.
 CC Pfam; PF00009; GTP_EFTU; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFCT.
 CC TIGRFAMS; TIGR01393; lepa; 1.
 CC TIGRFAMS; TIGR00231; small GTP; 1.
 CC PROSITE; PS00301; EFATOR_GTP; 1.
 CC GTP-binding; Complete proteome.
 CC NP_BIND 51 58 GTP (BY SIMILARITY).
 CC NP_BIND 120 124 GTP (BY SIMILARITY).
 CC NP_BIND 174 177 GTP (BY SIMILARITY).
 CC SEQUENCE 646 AA; 71329 MW; 256DC1EAB894C4A3 CRC64;

Query Match 45.8%; Score 44; DB 1; Length 646;
 Best Local Similarity 71.4%; Pred. No. 21;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KINLADRLMLGSGV 14
 DB 57 KSTLADRLMLQTV 70

RESULT 14

ID LEPA MYCTU STANDARD; PRT; 653 AA.
 AC P71739;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GTP-binding protein lepa.
 GN LEPA OR RV2404C OR MT2476 OR MTCX253.16 OR MB2427C.
 OS Mycobacterium tuberculosis, and
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773, 1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=H37RV;
 RX MEDLINE=98295997; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Ketchum J.A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.bovis; STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duchoy S., Grondin S., Lacroix C., Mensepe C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.
 CC Lepa subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z81368; CAB03723.1; -
 CC EMBL; AE007086; AAK46772.1; -
 CC EMBL; BX248342; CAD97288.1; -
 CC PIR; G70683; G70683.
 CC HSP; P13551; 2EFG.
 CC TIGR; MT2476; -
 CC Tuberculist; RV2404C; -
 CC HAMAP; MF 00071; -
 CC InterPro; IPR000795; EF_GTPbind.
 CC InterPro; IPR000640; EFG_C.
 CC InterPro; IPR009022; EFG_III_V.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR006297; LepA.
 CC InterPro; IPR005225; Small GTP.
 CC InterPro; IPR009000; Translat_factor.
 CC Pfam; PF00679; EFG_C; 1.
 CC Pfam; PF00009; GTP_EFTU; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFCT.
 CC TIGRFAMS; TIGR01393; lepa; 1.
 CC TIGRFAMS; TIGR00231; small GTP; 1.
 CC PROSITE; PS00301; EFATOR_GTP; 1.
 CC GTP-binding; Complete proteome.
 CC NP_BIND 59 66 GTP (BY SIMILARITY).
 CC NP_BIND 124 128 GTP (BY SIMILARITY).
 CC NP_BIND 178 181 GTP (BY SIMILARITY).
 CC SEQUENCE 653 AA; 72395 MW; DA4AFE10E6C25755 CRC64;

Query Match 45.8%; Score 44; DB 1; Length 653;
 Best Local Similarity 71.4%; Pred. No. 22;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KINLADRLMLGSGV 14

Search completed: March 15, 2004, 14:01:00
Job time : 5.64706 secs

Db 65 KSTLADRLMLQTV 78

```
RESULT 15
HS11 CAEEL STANDARD; PRT; 145 AA.
AC P34656;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heat shock protein HSP16-1.
GN (HSP16-1A OR T27E4.2) AND (HSP16-1B OR T27E4.8).
OS Caenorhabditis elegans
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85295957; PubMed=4033652;
RA Russnak R.H.; Candido E.P.M.;
RT "Locus encoding a family of small heat shock genes in Caenorhabditis
RT elegans: two genes duplicated to form a 3.8-kilobase inverted
RT repeat.";
RL Mcl. Cell. Biol. 5:1268-1278(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Bradshaw H.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 41-63 FROM N.A.
RX MEDLINE=87231065; PubMed=3588308;
RA Kay R.J.; Russnak R.H.; Jones D.; Mathias C.; Candido E.P.M.;
RT "Expression of intron-containing C. elegans heat shock genes in mouse
RT cells demonstrates divergence of 3' splice site recognition sequences
RT between nematodes and vertebrates, and an inhibitory effect of heat
RT shock on the mammalian splicing apparatus.";
RL Nucleic Acids Res. 15:3723-3741(1987).
CC -! SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K03273; AAA28068.1; -
CC EMBL; U64837; AAB04842.1; -
CC EMBL; U64837; AAB04839.1; -
CC PIR; B24289; B24289.
CC WormPep; T27E4.2; CE14249.
CC WormPep; T27E4.8; CE14249.
CC InterPro; IPR001436; Crystallin_alpha.
CC InterPro; IPR002068; Hsp20.
CC InterPro; IPR008978; HSP20_chap.
CC Pfam; PF00011; HSP20; 1.
CC PRINTS; PR00299; ACRYSTALLIN.
CC PROSITE; PS01031; HSP20; 1.
CC Heat shock; Multigene family.
CC SEQUENCE 145 AA; 16253 MW; 06C36A1F06D15A11 CRC64;
```

Query Match 44.8%; Score 43; DB 1; Length 145;
Best Local Similarity 45.0%; Pred. No. 6.6;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 KINLADRLMLGLSGVQEIKEQ 20
Db 65 KINLDGHTLSIQGEQLKTE 84

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:49:12 ; Search time 29.1765 seconds
(without alignments)
216.283 Million cell updates/sec

Title: US-09-171-432A-46
Perfect score: 96
Sequence: 1 KINLADRLGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	251	12 Q9ENP8	Q9enp8 hepatitis a
2	96	100.0	251	12 Q9ENN9	Q9enn9 hepatitis a
3	96	100.0	251	12 Q9ENQ4	Q9eng4 hepatitis a
4	96	100.0	251	12 Q9ENQ2	Q9eng2 hepatitis a
5	96	100.0	251	12 Q9ENN2	Q9enn2 hepatitis a
6	96	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
7	96	100.0	251	12 Q9ENQ6	Q9eng6 hepatitis a
8	96	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
9	96	100.0	251	12 Q9ENP4	Q9enp4 hepatitis a
10	96	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
11	96	100.0	251	12 Q9ENNA	Q9enn4 hepatitis a
12	96	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
13	96	100.0	251	12 Q9ENP3	Q9enp3 hepatitis a
14	96	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
15	96	100.0	251	12 Q9ENQ5	Q9eng5 hepatitis a
16	96	100.0	251	12 Q9ENQ5	Q9eng5 hepatitis a

17	96	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
18	96	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
19	96	100.0	251	12 Q9ENP6	Q9enp6 hepatitis a
20	96	100.0	251	12 Q9ENQ8	Q9eng8 hepatitis a
21	96	100.0	251	12 Q9ENP3	Q9enp3 hepatitis a
22	96	100.0	251	12 Q9ENQ0	Q9eng0 hepatitis a
23	96	100.0	1124	12 Q84780	Q84780 hepatitis a
24	96	100.0	1161	12 Q05794	Q05794 hepatitis a
25	96	100.0	2216	12 Q98MA2	Q98ma2 hepatitis a
26	96	100.0	2218	12 Q67824	Q67824 hepatitis a
27	96	100.0	2218	12 Q67817	Q67817 hepatitis a
28	96	100.0	2225	12 Q9DL32	Q9dl32 hepatitis a
29	96	100.0	2227	12 Q9WMA0	Q9wma0 hepatitis a
30	96	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
31	96	100.0	2227	12 Q67825	Q67825 hepatitis a
32	96	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
33	96	100.0	2227	12 Q8QV03	Q8qv03 hepatitis a
34	96	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
35	96	100.0	2227	12 Q67826	Q67826 hepatitis a
36	96	100.0	2227	12 Q8VON6	Q8von6 hepatitis a
37	96	100.0	2227	12 Q91FH5	Q91fh5 hepatitis a
38	96	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
39	92	95.8	251	12 Q9ENP0	Q9enp0 hepatitis a
40	92	95.8	251	12 Q9ENN5	Q9enn5 hepatitis a
41	92	95.8	251	12 Q9ENN7	Q9enn7 hepatitis a
42	91	94.8	2225	12 Q9DWR1	Q9dwr1 hepatitis a
43	91	94.8	2227	12 Q8QRI6	Q8qri6 hepatitis a
44	88	91.7	251	12 Q9ENR0	Q9enr0 hepatitis a
45	88	91.7	251	12 Q9ENN8	Q9enn8 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENP8	PRELIMINARY;	PRT;	251 AA.
ID Q9ENP8			
AC Q9ENP8			
DT 01-MAR-2001	(TRENBLrel. 16, Created)		
DT 01-MAR-2001	(TRENBLrel. 16, Last sequence update)		
DE 01-MAR-2001	(TRENBLrel. 16, Last annotation update)		
DE Polyprotein (Fragment)			
OS Hepatitis A virus.			
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC Hepatovirus.			
OX NCBI_TaxID=12092;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=A304;			
RA Fujiwara K.;			
RT "hepatitis A virus";			
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AB047685; BAB12173.1; -			
FT NON TER	1	251	
SQ SEQUENCE	251 AA;	28693 MW;	C32AD5651506751A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20

Db 125 KINLADRLGLSGVQEIKEQ 144

RESULT 2

ID Q9ENN9	PRELIMINARY;	PRT;	251 AA.
AC Q9ENN9			
DT 01-MAR-2001	(TRENBLrel. 16, Created)		
DT 01-MAR-2001	(TRENBLrel. 16, Last sequence update)		

DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A68;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047674; BAB12182.1; -.
FT NON_TER 1 251
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
DB 125 KINLADRMGLSGVQEIKEQ 144

RESULT 3
Q9ENQ4 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ4;
AC Q9ENQ4;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A201;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047659; BAB12167.1; -.
FT NON_TER 1 251
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
DB 125 KINLADRMGLSGVQEIKEQ 144

RESULT 4
Q9ENQ2 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ2;
AC Q9ENQ2;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A205;
RA Fujiwara K.;

RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047661; BAB12169.1; -.
FT NON_TER 1 251
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28723 MW; 558A2D3664C7343C CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
DB 125 KINLADRMGLSGVQEIKEQ 144

RESULT 5
Q9ENN2 PRELIMINARY; PRT; 251 AA.
ID Q9ENN2;
AC Q9ENN2;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A9;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047681; BAB12189.1; -.
FT NON_TER 1 251
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
DB 125 KINLADRMGLSGVQEIKEQ 144

RESULT 6
Q9ENP2 PRELIMINARY; PRT; 251 AA.
ID Q9ENP2;
AC Q9ENP2;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=A5;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON_TER 1 251
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
125 KINLADRMGLSGVQEIKEQ 144

Db

RESULT 7

Q9ENQ6 PRELIMINARY; PRT; 251 AA.

AC Q9ENQ6
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A162;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
125 KINLADRMGLSGVQEIKEQ 144

Db

RESULT 8

Q9ENR1 PRELIMINARY; PRT; 251 AA.

AC Q9ENR1
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A1;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1; -
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
125 KINLADRMGLSGVQEIKEQ 144

Db

RESULT 9

Q9ENP1 PRELIMINARY; PRT; 251 AA.

AC Q9ENP1
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A503;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; -
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
125 KINLADRMGLSGVQEIKEQ 144

Db

ID Q9ENP1 PRELIMINARY; PRT; 251 AA.

AC Q9ENP1
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A503;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; -
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
125 KINLADRMGLSGVQEIKEQ 144

Db

RESULT 10

Q9ENP4 PRELIMINARY; PRT; 251 AA.

AC Q9ENP4
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A414;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047669; BAB12177.1; -
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||
125 KINLADRMGLSGVQEIKEQ 144

Db

RESULT 11

Q9ENQ9 PRELIMINARY; PRT; 251 AA.

AC Q9ENQ9
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

```

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A159;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 125 KINLADRMGLSGVQEIKEQ 144
|||||
|||||

RESULT 12
Q9ENNA PRELIMINARY; PRT; 251 AA.
ID Q9ENNA4
AC Q9ENNA4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A77;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB047679; BAB12187.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 125 KINLADRMGLSGVQEIKEQ 144
|||||
|||||

RESULT 13
Q9ENP5 PRELIMINARY; PRT; 251 AA.
ID Q9ENP5
AC Q9ENP5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 125 KINLADRMGLSGVQEIKEQ 144
|||||
|||||

RESULT 14
Q9ENQ3 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ3
AC Q9ENQ3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A204;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB047660; BAB12168.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 125 KINLADRMGLSGVQEIKEQ 144
|||||
|||||

RESULT 15
Q9ENP7 PRELIMINARY; PRT; 251 AA.
ID Q9ENP7
AC Q9ENP7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A306;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 125 KINLADRMGLSGVQEIKEQ 144
|||||
|||||

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
Db 125 KINLADRMGLSGVQEIKEQ 144
|||||
|||||
```

Db 125 KINLADRLGLSGVQEIKEQ 144

Search completed: March 15, 2004, 14:05:22
Job time : 29.1765 secs

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:24:17 ; Search time 55.2941 Seconds
(without alignments)
127.748 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesep29Jan04:*
1: Genesep29Jan04:*
2: Genesep29Jan04:*
3: Genesep29Jan04:*
4: Genesep29Jan04:*
5: Genesep29Jan04:*
6: Genesep29Jan04:*
7: Genesep29Jan04:*
8: Genesep29Jan04:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	129	100.0	25	AAW42930 Immunogen
2	129	100.0	26	AAW42930 Immunogen
3	129	100.0	366	AAW42930 Immunogen
4	129	100.0	854	AAW42930 Immunogen
5	129	100.0	993	AAW42930 Immunogen
6	129	100.0	993	AAW42930 Immunogen
7	129	100.0	1077	AAW42930 Immunogen
8	129	100.0	1091	AAW42930 Immunogen
9	129	100.0	2227	AAW42930 Immunogen
10	129	100.0	2227	AAW42930 Immunogen
11	129	100.0	2227	AAW42930 Immunogen
12	129	100.0	2227	AAW42930 Immunogen
13	129	100.0	2227	AAW42930 Immunogen
14	129	100.0	2227	AAW42930 Immunogen
15	129	100.0	2227	AAW42930 Immunogen
16	129	100.0	2227	AAW42930 Immunogen
17	129	100.0	2227	AAW42930 Immunogen
18	129	100.0	2227	AAW42930 Immunogen
19	124	96.1	2227	AAW42930 Immunogen
20	124	96.1	2227	AAW42930 Immunogen
21	124	96.1	2227	AAW42930 Immunogen
22	124	96.1	2227	AAW42930 Immunogen
23	117	90.7	839	AAW42930 Immunogen
24	103	79.8	20	AAW42930 Immunogen
25	103	79.8	21	AAW42930 Immunogen

26	65	50.4	20	2	AAW42925	AAW42925 Immunogen
27	65	50.4	21	4	AAW42925	AAW42925 Immunogen
28	51	39.5	150	4	AB668507	AB668507 Drosophil
29	49	38.0	420	4	AB668507	AB668507 Drosophil
30	49	38.0	444	2	AAW20824	AAW20824 H. pylori
31	49	38.0	686	6	ADA33016	ADA33016 Acinetoba
32	48.5	37.6	2000	6	ABR52734	ABR52734 Protein s
33	48	37.2	187	3	ABR52734	ABR52734 Eucalyptu
34	48	37.2	221	4	ABR52734	ABR52734 Eucalyptu
35	48	37.2	523	7	ADCO1554	ADCO1554 Enterohae
36	48	37.2	1176	4	ABR59892	ABR59892 Drosophil
37	47.5	36.8	86	3	AAO26501	AAO26501 Human gly
38	47.5	36.8	241	7	ADB64347	ADB64347 Human pro
39	47.5	36.8	468	7	ADB64347	ADB64347 Human pro
40	47.5	36.8	530	3	AA342529	AA342529 Human ORF
41	47.5	36.8	530	5	ABP68940	ABP68940 Human pol
42	47.5	36.8	530	5	AAO26501	AAO26501 Human gly
43	46.5	36.0	23	2	AAW42930	AAW42930 Murine os
44	46.5	36.0	137	2	AAW42930	AAW42930 Murine os
45	46.5	36.0	139	2	AAW42930	AAW42930 Mature mu

ALIGNMENTS

RESULT 1
AAW42930
ID AAW42930 standard; peptide; 25 AA.
XX AAW42930;
XX AC
XX DT 28-APR-1998 (first entry)
XX DE
XX DE Immunogenic Hepatitis A virus peptide YK-1665.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX KW antibody.
XX OS Synthetic.
XX OS Hepatitis A virus.
XX PN WO9740147-A1.
XX PD 30-OCT-1997.
XX PF 18-APR-1997; 97WO-US006891.
XX PR 19-APR-1996; 96US-0015644P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudiyakov YE;
XX DR WPI; 1997-535831/49.
XX PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal.
XX PS Claim 18; Page 112; 140pp; English.
XX CC Peptides AAW42922-30 are immunogenic peptides corresponding to
XX CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
XX CC substantially similar to a portion of the amino acid sequence of the P2A
XX CC protein of HAV corresponding to amino acids 792-980. Compositions to HAV
XX CC containing the peptides can be used to induce an immune response to HAV
XX CC in a mammal. The peptides can also be used to detect the presence of
XX CC antibodies against HAV in mammalian serum. The peptides can also be used
XX CC to make an antibody against HAV by administering the peptide to a mammal
XX SQ Sequence 25 AA;

Query Match 100.0%; Score 129; DB 2; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Dy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

RESULT 2
AAB69447
ID AAB69447 standard; peptide; 26 AA.
XX AAB69447;
XX 20-APR-2001 (first entry)
XX Synthetic HAV P2A peptide, SEQ ID NO: 47.
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
XX antigen; major structural capsid polypeptide; HAV antibody detection.
XX Hepatitis A virus.
XX Synthetic.
XX WO200105824-A2.
XX 25-JAN-2001.
XX 14-JUL-2000; 2000WO-US019267.
XX 15-JUL-1999; 99US-0144412P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Fields HA, Khudyakov YE;
XX WPI; 2001-112681/12.
XX Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines.
XX Claim 13; Page 98; 130pp; English.

The present sequence is one of a number of synthetic peptides which are immunoreactive with hepatitis A virus (HAV) antibodies. The peptides comprise antigenic epitopes of the major structural capsid polypeptides or non-structural polypeptides of HAV with one or more glutamine molecules at the carboxy end of the peptide. The peptides are used to detect the presence of antibodies against HAV in mammalian serum. to detect the presence of HAV in a human or animal through the binding of the peptide to an antibody, to detect acute phase infection by detecting IGM antibodies in mammalian serum and detecting convalescence in a mammal. The peptides are used to detect or quantify HAV antibodies in samples in clinical or research-based assays using immunoblotting, fluorescence in situ hybridisation analysis, gel-mobility shift assays, tracking of radioactive or bioluminescent markers, chromatography or electrophoresis. The peptides are used to induce an immune response to HAV when administered to a human or animal. Glutamine at the carboxy end of the peptides enhances the IGM antibody reactivity

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Dy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

RESULT 3
AAP50230

Query Match 100.0%; Score 129; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Dy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

RESULT 3
AAP50230

ID AAP50230 standard; protein; 366 AA.
XX AAP50230;
XX 28-NOV-1991 (first entry)
XX Sequence of hepatitis A virus (HAV) surface protein (VP-1).
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
XX diagnostic assay.
XX Hepatitis A virus.
XX EPI38704-A.
XX 24-APR-1985.
XX 09-OCT-1984; 84EP-00402025.
XX 14-OCT-1983; 83US-00541836.
XX 02-MAR-1984; 84US-00585942.
XX (MERI) MERCK & CO INC.
XX Hughes JV, Scolnick EM, Tomassini JE;
XX WPI; 1985-100818/17.
XX N-P8DB; AAN50274.
XX New hepatitis A virus surface protein - useful for binding to
XX neutralising antibodies to the virus.
XX Claim 21; Page 46-48; 49pp; English.
XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
XX surfactant and a reducing agent. The viral proteins are sepd. and the
XX protein of molecular wt. 33000 daltons is sepd
XX Sequence 366 AA;

Query Match 100.0%; Score 129; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 2.4e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Dy 319 QRLKYAQEELSNEVLPPPRKMGFLF 343

RESULT 4
AAP50287
ID AAP50287 standard; protein; 854 AA.
XX AAP50287;
XX 25-MAR-2003 (revised)
XX 30-NOV-1991 (first entry)
XX Hepatitis A virus (HAV) peptide corresponding to the capsid protein
XX region of poliovirus RNA.
XX Hepatitis A virus assay; antigen; antibody.
XX Hepatitis A virus.
XX WO9501517-A.
XX 11-APR-1985.
XX 27-SEP-1984; 84WO-US001552.
XX 30-SEP-1983; 83US-00537911.
XX

PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
PI Racanelli VR;
XX WPI; 1985-098846/16.
DR N-PSDB; AAN50330.
XX
PT New hepatitis A virus CDNA - useful in assays for the virus and for
PT prodn. of the viral antigen and antibodies to it.
XX
XX Example; Fig 7; 60pp; English.
XX
CC The inventors claim HAV CDNA and a method for producing it, whereby large
CC ams. can be obt'd. economically. The CDNA is useful in the assay for
CC detection of HAV quickly and easily and with high sensitivity and
CC specificity. The HAV CDNA is also used in the prodn. of HAV antigen or
CC antibodies to it. The antibodies may be monoclonal. (Updated on 25-MAR-
CC 2003 to correct PA field.)
XX
SQ Sequence 854 AA;
Query Match 100.0%; Score 129; DB 1; Length 854;
Best Local Similarity 100.0%; Pred. No. 5.9e-11; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMGFLF 834
RESULT 5
AAP50116
ID AAP50116 standard; protein; 993 AA.
XX AAP50116;
XX
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-SEP-1991 (first entry)
XX
DE Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
DE and VP-4.
XX
XX Antigenic protein; immunogen; vaccine.
XX
OS Hepatitis A virus; (strain CR326).
XX
XX EP154587-A.
XX
PD 11-SEP-1985.
XX
XX 27-FEB-1985; 85EP-00400369.
XX
XX 02-MAR-1984; 84US-00585818.
XX
XX (MERI) MERCK & CO INC.
XX
XX Lineneyer DL, Menke JG, Rueben RG, Mitra SW;
XX WPI; 1985-224964/37.
DR N-PSDB; AAN50139.
XX
XX New nucleotide sequences coding for hepatitis A virus antigens - useful
PT for eliciting normal immune response and in vaccines for protecting
PT against the virus.
XX
XX Example; Page 11-17; 32pp; English.
PS
XX Within the sequence in AAN50139 is encoded the information necessary to
CC make the antigenic proteins of HAV. The sequences encoding for the
CC structural proteins begin at base 403. The key sub-unit sequences within
CC VP-1, designated Sequences I,II,III,IV, and V, start, respectively at

CC 1882, 1963, 1999, 2146, 2347. Other nucleotide sequences which are
CC valuable as encoding antigenic proteins are the sequences from base 1749
CC to base 2722; from base 1487 to base 2980 and from base 1644 to base
CC 2722. The sequence from base 1749 to base 2722 is esp. valuable as a
CC vector for producing antigen protein. Sequences II-V are claimed. X in
CC AAP50116 denotes the translation of a stop codon. (Updated on 25-MAR-2003
CC to correct PI field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 993 AA;
Query Match 100.0%; Score 129; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 7e-11; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 946 QRLKYAQEELSNEVLPPPRKMGFLF 970
RESULT 6
AAP50231
ID AAP50231 standard; protein; 993 AA.
XX AAP50231;
AC AAP50231;
XX
DT 28-NOV-1991 (first entry)
XX
DE Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).
XX
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
XX diagnostic assay.
XX
OS Hepatitis A virus.
XX
XX Key Location/Qualifiers
FH Protein 628..993
FT /note= "claimed; X denotes translated stop codons and
FT unspecified triplets"
XX
XX EP138704-A.
XX
XX 24-APR-1985.
XX
XX 09-OCT-1984; 84EP-00402025.
XX
XX 14-OCT-1983; 83US-00541836.
PR 02-MAR-1984; 84US-00585942.
XX
XX (MERI) MERCK & CO INC.
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
PI WPI; 1985-100818/17.
DR N-PSDB; AAN50274.
XX
XX New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus.
XX
XX Disclosure; Page 17-23; 49pp; English.
XX
XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
CC surfactant and a reducing agent. The viral proteins are sepd. and the
CC protein of molecular wt. 33000 daltons is sepd
XX
SQ Sequence 993 AA;
Query Match 100.0%; Score 129; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 7e-11; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

KW	HAV HM-175; chronic liver disease; picornavirus.
XX	Hepatitis A virus.
XX	
XX	Location/Qualifiers
FT	1..711
FT	/note= "X's correspond to nonsense codons, i.e. this
FT	region is not an ORF"
FT	238..1091
FT	/label= ORF
FT	/note= "second putative initiation codon at position 240"
FT	
FN	USN7789262-N.
XX	
XX	15-DEC-1992.
XX	
XX	06-NOV-1991; 91US-00789262.
XX	
XX	30-SEP-1983; 88US-00536911.
PR	27-SEP-1984; 84US-00654942.
PR	06-OCT-1988; 88US-00256135.
XX	
XX	(USSH) US DEPT HEALTH & HUMAN SERVICE.
PA	
PI	Ticahurst JR, Baltimore D, Feinstone SM, Purcell RH;
PI	Racaniello VR, Baroudy BM, Emerson SU;
XX	
DR	WPI: 1993-067429/08.
DR	N-PSDB; AAQ36934.
XX	
PT	Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of
PT	antigen and antibodies.
XX	
PS	Disclosure; Fig 7; 65pp; English.
XX	
CC	HAV virion RNA was extracted from the livers of marmosets which had been
CC	inoculated with HAV (the HAV had previously been passaged twice in
CC	marmosets). The RNA was used to prepare ds cDNA clones by standard
CC	methods. Clones contg. inserts which hybridised to RNA from HAV-infected
CC	African Green Monkey Kidney cells were selected for further analysis. A
CC	5.4kb restriction map (about 9% of the HAV genome) was constructed from
CC	5'-terminus was determined. The sequence of the first 3.3kb (approx.) from the
CC	entire clone and an open reading frame was identified starting at
CC	position 238. A comparison of the predicted HAV amino acid sequences with
CC	the known capsid protein sequences of other picornaviruses (poliovirus,
CC	foot and mouth disease virus and encephalomyelitis virus) revealed areas
CC	of local homology. (Note: Revised entry submitted to correct the patent
CC	number format of US Government-owned NTIS applications to prevent clashes
CC	with ongoing US granted patent numbers. For further information please
CC	visit the Derwent web site at www.derwent.com/api/updatee/ntis us.html.)
CC	(Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to
XX	correct OS field.)
SQ	Sequence 1091 AA;
	Query Match 100.0%; Score 129; DB 2; Length 1091;
	Best Local Similarity 100.0%; Pred.No. 7.7e-11;
	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ORLKYAQEELSNEVLPPPRKMGFLF 25
DB	1047 QRLKYAQEELSNEVLPPPRKMGFLF 1071
RESULT 9	
AAR05697	
ID	AAR05697 standard; protein; 2227 AA.
XX	
AC	AAR05697;
XX	
DT	24-OCT-2003 (revised)
DT	27-AUG-2003 (revised)

```

Db          946 QRLKYAQEELSNEVLPPPRKMKGLF 970

RESULT 7
AAW95559
ID  AAW95559 standard; protein; 1077 AA.
XX
XX  AAW95559;
XX
XX  28-APR-1999 (first entry)
XX
XX  A partial hepatitis A virus (HAV) protein.
DE
XX  Hepatitis A virus protein; HAV; P2 region;
XX  cell-culture-adapted HAV strain; infection; accelerated growth.
XX  Hepatitis A virus.
OS
XX
XX  US5849562-A.
PN
XX
XX  15-DEC-1998.
XX
XX  06-JUN-1995; 95US-00468926.
XX
XX  30-SEP-1983; 83US-00537911.
XX  27-SEP-1984; 84US-00654942.
XX  06-OCT-1988; 88US-00256135.
XX  06-NOV-1991; 91US-00788262.
XX
XX  (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX  Emerson SU, Purcell RH;
XX
XX  WPI: 1999-094412/08.
XX  N-PSDB; AAX01006.
XX
XX  Chimeric hepatitis A virus strains - with P2 region from cell-culture-
XX  adapted strain in wild-type genome.
XX
XX  Disclosure; Fig 7A-L; 36pp; English.
XX
XX  The present sequence represents a partial hepatitis A virus (HAV)
XX  protein. The specification describes a DNA construct consisting of a wild
XX  -type HAV genome in which the P2 region is replaced by the P2 region from
XX  a cell-culture-adapted HAV strain. The construct is used to demonstrate
XX  that mutations in the P2 region of a cell-culture-adapted HAV strain are
XX  sufficient for establishment of infection and accelerated growth in cell
XX  culture
XX
XX  Sequence 1077 AA;
SQ
Query Match          100.0%; Score 129; DB 2; Length 1077;
Best Local Similarity 100.0%; Pred. NO. 7.6e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  QRLKYAQEELSNEVLPPPRKMKGLF 25
        |||||||
Db      1033 QRLKYAQEELSNEVLPPPRKMKGLF 1057

RESULT 8
AAR32426
ID  AAR32426 standard; protein; 1091 AA.
XX
XX  AAR32426;
XX
XX  27-AUG-2003 (revised)
XX  25-MAR-2003 (revised)
XX  17-DEC-2001 (revised)
XX  10-JUN-1993 (first entry)
XX
XX  Translated from 5' region of Hepatitis A Virus genomic clone.
DE
XX

```

DT 25-MAR-2003 (revised)
DT 15-AUG-1990 (first entry)
XX
XX Attenuated hepatitis A virus.
XX
XX Hepatitis A virus; vaccine; attenuated.
KW Hepatitis A virus; strain HM-175.
OS
XX
XX
XX Key Location/Qualifiers
FH Region 1. .23
FT /label= VP4 = 1A
FT Region 24. .245
FT /label= VP2 = 1B
FT Region 246. .491
FT /label= VP3 = 1C
FT Region 492. .791
FT /label= VP1 = 1D
FT Region 792. .980
FT /label= 2A
FT Region 981. .1087
FT /label= 2B
FT Region 1088. .1422
FT /label= 2C
FT Region 1423. .1496
FT /label= 3A
FT Region 1497. .1519
FT /label= 3B = VPg
FT Region 1520. .1738
FT /label= 3C
FT Region 1739. .2227
FT /label= 3D
XX
XX US4894228-A.
PN
XX
XX 16-JAN-1990.
PD
XX
XX 12-JUL-1988; 88US-00217824.
PF
XX
XX 19-SEP-1984; 84US-00652067.
PR
XX 09-SEP-1986; 86US-00905146.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
PA
XX
XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
PI Daemer RJ, Gust ID;
PI
XX WPI; 1990-075557/10.
DR N-PSDB; AAQ03512.
DR
XX Vaccine against hepatitis A virus infection - comprises novel attenuated
PT hepatitis A virus strain.
PT
XX Claim 1; Fig 1; 18pp; English.
PS
XX The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 129; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORLKYAQEELSNEVLPPPRMKGLF 25
|||||

Db 810 ORLKYAQEELSNEVLPPPRMKGLF 834
RESULT 10
AAW34074
ID AAW34074 standard; protein; 2227 AA.
XX
XX AAW34074;
XX
XX 17-OCT-2003 (revised)
DT 27-APR-1998 (first entry)
XX
XX Hepatitis A virus HM-175 protein sequence.
XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
KW vaccine.
XX
XX Hepatitis A virus; HM-175.
OS
XX
XX Key Location/Qualifiers
FH Protein 1. .23
FT /label= VP4
FT Protein 24. .245
FT /label= VP2
FT Protein 246. .491
FT /label= VP3
FT Protein 492. .791
FT /label= VP1
FT Protein 792. .980
FT /label= 2A
FT Protein 981. .1087
FT /label= 2B
FT Protein 1088. .1422
FT /label= 2C
FT Protein 1423. .1496
FT /label= 3A
FT Protein 1497. .1519
FT /label= 3B
FT Protein 1520. .1738
FT /label= 3C
FT Protein 1739. .2227
FT /label= 3D
XX
XX WO9740166-A2.
PN
XX
XX 30-OCT-1997.
PD
XX
XX 18-APR-1997; 97WO-US006506.
PF
XX
XX 19-APR-1996; 96US-0015642P.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Raychaudhuri G, Emerson SU, Purcell RH;
PI
XX WPI; 1997-535850/49.
DR N-PSDB; AAT93023.
DR
XX Human attenuated HAV genome containing simian HAV 2C gene - useful as
PT vaccines against HAV infection.
PT
XX Disclosure; Fig 13A-D; 66pp; English.
PS
XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
CC 175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
CC by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
CC construct (1) comprises a genome of HAV, where the genome is a human
CC attenuated HAV genome in which a region of the 2C gene has been replaced
CC by a corresponding region from a 2C gene of a simian AGM-27 HAV genome
CC (see AAT93024). The region of the 2C gene from AGM-27 contained in the
CC construct preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript
CC of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3)

21-FEB-2002.

15-AUG-2001; 2001WO-IB001808.

17-AUG-2000; 2000US-0225767P.

29-AUG-2000; 2000US-0229175P.

03-NOV-2000; 2000US-00705547.

(TRIP-) TRIPEP AB.

Sallberg M, Hultgren C;

WPI; 2002-241837/29.

N-PSDB; AAD31766.

Vaccine compositions for treating and preventing disease, preferably hepatitis C virus infection, comprises ribavirin and antigen that has epitope present in hepatitis C virus.

Claim 11; Page 82-87; 120pp; English.

The invention relates to a composition comprising ribavirin and an antigen preferably non structural 3 protein (NS3)/4A fragment of hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV sequence. The composition is useful for enhancing an immune response to a hepatitis C antigen in humans, domestic, sport or pet species and as vaccines for treating and preventing HCV infections. The composition is also useful for treating viral, bacterial, fungal diseases and cancer. The present sequence is hepatitis A virus (HAV) protein

Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 5; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.7e-10;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGKLP 25

Db 810 QRLKYAQEELSNEVLPPPRKMGKLP 834

RESULT 14

ABG31729

ID ABG31729 standard; protein; 2227 AA.

AC ABG31729;

XX

XX

29-AUG-2003 (revised)

DT 29-NOV-2002 (first entry)

XX

DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.

XX

XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide; HAV 4380.

KW

XX

XX Hepatitis A virus; strain HM-175.

OS

PN US6423318-B1.

XX

XX

23-JUL-2002.

XX

31-AUG-2000; 2000US-00653499.

XX

17-SEP-1993; 93WO-US008610.

PR

17-APR-1995; 95US-00397232.

PR

07-JUN-1995; 95US-00475886.

XX

(USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX

PA Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

PI WPI: 2002-680946/73.

XX

DR

Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRLKYAQEELSNEVLPPPRXMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRXMKGLF 834

Search completed: March 15, 2004, 14:00:00
Job time : 55.2941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:15:47 ; Search time 14.1176 Seconds
(without alignments)
91.421 Million cell updates/sec

Title: US-09-171-432A-47
Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	1091	6	Patent No. 5516630
2	129	100.0	2227	3	US-08-475-886-2
3	129	100.0	2227	3	US-08-475-886-6
4	129	100.0	2227	3	US-08-397-232-2
5	129	100.0	2227	3	US-08-397-232-4
6	129	100.0	2227	3	US-09-171-387-2
7	129	100.0	2227	4	US-09-653-499-2
8	129	100.0	2227	4	US-09-653-499-6
9	129	100.0	2227	4	US-10-104-966-12
10	129	100.0	2227	4	US-10-135-988-2
11	129	100.0	2227	4	US-10-135-988-6
12	124	96.1	2227	3	US-08-475-886-4
13	124	96.1	2227	4	US-09-653-499-4
14	124	96.1	2227	4	US-10-135-988-4
15	117	90.7	839	1	US-08-087-016-2
16	49	38.0	686	4	US-09-328-352-4303
17	47	36.4	180	4	US-09-252-991A-18339
18	46.5	36.0	23	1	US-07-800-364B-4
19	46.5	36.0	23	5	PCT-US91-03388-4
20	46.5	36.0	139	1	US-08-278-729A-8
21	46.5	36.0	139	1	US-08-155-343A-8
22	46.5	36.0	139	1	US-08-406-672-8
23	46.5	36.0	139	1	US-08-643-563A-8
24	46.5	36.0	139	1	US-08-643-763A-8
25	46.5	36.0	139	1	US-08-462-623-8
26	46.5	36.0	139	1	US-08-451-953A-8
27	46.5	36.0	139	2	US-08-445-468A-8

```
28 46.5 36.0 139 2 US-08-461-397A-8 Sequence 8, Appli
29 46.5 36.0 139 2 US-08-912-088-8 Sequence 8, Appli
30 46.5 36.0 139 3 US-08-278-730A-8 Sequence 8, Appli
31 46.5 36.0 139 3 US-08-445-467-8 Sequence 8, Appli
32 46.5 36.0 139 3 US-08-480-515A-8 Sequence 8, Appli
33 46.5 36.0 139 3 US-08-414-033A-8 Sequence 8, Appli
34 46.5 36.0 139 3 US-08-271-556A-6 Sequence 6, Appli
35 46.5 36.0 139 3 US-08-440-894A-8 Sequence 8, Appli
36 46.5 36.0 139 4 US-09-170-936-8 Sequence 8, Appli
37 46.5 36.0 139 4 US-08-461-113-8 Sequence 8, Appli
38 46.5 36.0 139 4 US-08-456-033-8 Sequence 7, Appli
39 46.5 36.0 139 4 US-08-643-321-7 Sequence 8, Appli
40 46.5 36.0 139 4 US-09-464-206-8 Sequence 8, Appli
41 46.5 36.0 139 4 US-08-404-113A-8 Sequence 8, Appli
42 46.5 36.0 139 5 PCT-US92-01968-8 Sequence 8, Appli
43 46.5 36.0 139 5 PCT-US93-07190-8 Sequence 8, Appli
44 46.5 36.0 139 5 PCT-US93-07231-8 Sequence 8, Appli
45 46.5 36.0 139 5 PCT-US93-08742-8 Sequence 8, Appli
```

ALIGNMENTS

```
RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO:2
; LENGTH: 1091
5516630-2
```

```
Query Match 100.0%; Score 129; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 1047 QRLKYAQEELSNEVLPPPRKMGFLF 1071
```

```
RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
```

```
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
   |||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 3
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
   |||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 4
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
   |||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 5
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
   |||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 6
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
```


; APPLICATION NUMBER: US60/015,642
 ; FILING DATE: 19-APR-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: William S. Feiler
 ; REGISTRATION NUMBER: 26,728
 ; REFERENCE/DOCKET NUMBER: 2026-4229US1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; INFORMATION FOR SEQ ID NO: 2
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2227 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-171-387-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 7
 US-09-653-499-2
 ; Sequence 2, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 ; FILE REFERENCE: 20264262US2
 ; CURRENT APPLICATION NUMBER: US/09/653,499
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18
 ; PRIOR APPLICATION NUMBER: 08/397,232
 ; PRIOR FILING DATE: 1995-03-10
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
 US-09-653-499-2

Query Match 100.0%; Score 129; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8
 US-09-653-499-6
 ; Sequence 6, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

; FILE REFERENCE: 20264262US2
 ; CURRENT APPLICATION NUMBER: US/09/653,499
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18
 ; PRIOR APPLICATION NUMBER: 08/397,232
 ; PRIOR FILING DATE: 1995-03-10
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: Attenuated (4380) HAV, strain HM-175
 US-09-653-499-6

Query Match 100.0%; Score 129; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 9
 US-10-104-966-12
 ; Sequence 12, Application US/10104966
 ; Patent No. 6680059
 ; GENERAL INFORMATION:
 ; APPLICANT: Matti Sallberg
 ; APPLICANT: Catharina Hultgren
 ; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: TRIPEP-23AUSC1
 ; CURRENT APPLICATION NUMBER: US/10/104,966
 ; CURRENT FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: 09/705,547
 ; PRIOR FILING DATE: 2000-11-03
 ; PRIOR APPLICATION NUMBER: 60/229,175
 ; PRIOR FILING DATE: 2000-08-29
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hepatitis A virus sequence
 US-10-104-966-12

Query Match 100.0%; Score 129; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 10
 US-10-135-988-2
 ; Sequence 2, Application US/10135988
 ; Patent No. 6680060
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 ; FILE REFERENCE: 20264262US3
 ; CURRENT APPLICATION NUMBER: US/10/135,988

```
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match
Best Local Similarity 100.0%; Score 129; DB 4; Length 2227;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 11
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4390) HAV, strain HM-175
US-10-135-988-6

Query Match
Best Local Similarity 100.0%; Score 129; DB 4; Length 2227;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 12
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match
Best Local Similarity 96.1%; Score 124; DB 3; Length 2227;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 13
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match
Best Local Similarity 96.1%; Score 124; DB 3; Length 2227;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 14
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
```

; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 96.1%; Score 124; DB 4; Length 2227;
Best Local Similarity 96.0%; Pred. No. 1e-10; Indels 0; Gaps 0;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKVAQEELSNEVLPPPRKMKGLF 25
|||:|||||:|||||:|||||
Db 810 ORLKVAQEELSNEVLPPPRKMKGLF 834

RESULT 15
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match 90.7%; Score 117; DB 1; Length 839;
Best Local Similarity 84.0%; Pred. No. 4.4e-10;
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ORLKVAQEELSNEVLPPPRKMKGLF 25
|||:|||||:|||||:|||||
Db 809 QREKYAREELSNEILPPPRKMKGLF 833

Search completed: March 15, 2004, 13:26:06
Job time : 14.1176 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:23:07 ; Search time 28.9706 Seconds
(without alignments)
182.213 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKYAQELSNEVLPPFRKMGKLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	25	US-09-171-432A-47	Sequence 47, Appl
2	129	100.0	352	US-10-272-459-45	Sequence 45, Appl
3	129	100.0	980	US-10-272-459-41	Sequence 41, Appl
4	129	100.0	2227	US-09-929-955-12	Sequence 12, Appl
5	129	100.0	2227	US-10-104-966-12	Sequence 12, Appl
6	129	100.0	2227	US-10-135-988-2	Sequence 2, Appl
7	129	100.0	2227	US-10-135-988-6	Sequence 6, Appl
8	125	96.9	836	US-10-272-459-40	Sequence 40, Appl
9	124	96.1	2227	US-10-135-988-4	Sequence 4, Appl
10	103	79.8	20	US-09-171-432A-41	Sequence 41, Appl
11	65	50.4	20	US-09-171-432A-42	Sequence 42, Appl
12	52	40.3	1193	US-10-369-493-5713	Sequence 5713, Ap
13	51	39.5	1240	US-10-369-493-4031	Sequence 4031, Ap
14	51	39.5	6642	US-10-369-493-5013	Sequence 5013, Ap
15	50	38.8	1168	US-10-369-493-3980	Sequence 3980, Ap

16	48	37.2	300	15	US-10-369-493-10595	Sequence 10595, A
17	47.5	36.8	241	15	US-10-104-047-2501	Sequence 2501, Ap
18	47.5	36.8	468	15	US-10-104-047-3508	Sequence 3508, Ap
19	47.5	36.8	530	15	US-10-108-260A-4331	Sequence 4331, Ap
20	47	36.4	592	15	US-10-369-493-1534	Sequence 1534, Ap
21	46.5	36.0	139	8	US-08-260-675-8	Sequence 8, Appli
22	46.5	36.0	139	14	US-10-050-050-8	Sequence 0, Appli
23	46.5	36.0	399	8	US-08-957-425-27	Sequence 27, Appl
24	46.5	36.0	399	8	US-08-260-675-23	Sequence 23, Appl
25	46.5	36.0	399	14	US-10-122-026-8	Sequence 8, Appli
26	46.5	36.0	399	14	US-10-050-050-23	Sequence 23, Appl
27	46.5	36.0	399	14	US-10-350-747-2	Sequence 2, Appli
28	46	35.7	263	9	US-09-864-761-37656	Sequence 37656, A
29	45.5	35.3	1596	10	US-09-909-567B-47	Sequence 47, Appl
30	45	34.9	20	10	US-09-171-432A-40	Sequence 40, Appl
31	45	34.9	223	9	US-09-895-913A-84	Sequence 84, Appl
32	45	34.9	445	14	US-10-177-293-296	Sequence 296, App
33	45	34.9	445	15	US-10-435-696-33	Sequence 33, Appl
34	45	34.9	756	16	US-10-389-566-1970	Sequence 1970, Ap
35	45	34.9	1323	16	US-10-389-566-1642	Sequence 1642, Ap
36	44.5	34.5	3562	15	US-10-341-434-109	Sequence 109, App
37	44	34.1	303	15	US-10-369-493-22337	Sequence 22337, A
38	44	34.1	541	15	US-10-291-172-670	Sequence 670, App
39	44	34.1	914	15	US-10-369-493-1851	Sequence 1851, Ap
40	43.5	33.7	222	14	US-10-156-761-11163	Sequence 11163, A
41	43.5	33.7	1114	10	US-09-840-743-14	Sequence 14, Appl
42	43	33.3	106	15	US-10-291-172-615	Sequence 615, App
43	43	33.3	171	15	US-10-094-749-2193	Sequence 2193, Ap
44	43	33.3	213	11	US-09-833-245-1852	Sequence 1852, Ap
45	43	33.3	223	14	US-10-029-386-34256	Sequence 34256, A

ALIGNMENTS

RESULT 1
US-09-171-432A-47
; Sequence 47, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171/432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1-25
; OTHER INFORMATION: /label=YK-1665
US-09-171-432A-47

Query Match      100.0%; Score 129; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

RESULT 2
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match      100.0%; Score 129; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 182 QRLKYAQEELSNEVLPPPRKMGFLF 206

RESULT 3
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match      100.0%; Score 129; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 4
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      100.0%; Score 129; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 5
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 129; DB 13; Length 2227;
```

```
Best Local Similarity 100.0%; Pred. No. 1e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 6
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227.
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 129; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 7
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 129; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8
US-10-272-459-40
; Sequence 40, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 94 kDa
US-10-272-459-40

Query Match 96.9%; Score 125; DB 14; Length 836;
Best Local Similarity 96.0%; Pred. No. 1.4e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 9
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 96.1%; Score 124; DB 13; Length 2227;
Best Local Similarity 96.0%; Pred. No. 5.5e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 10
US-09-171-432A-41
; Sequence 41, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
```

APPLICANT: Fields, Howard A.
APPLICANT: Khudiyakov, Yury E.
TITLE OF INVENTION: Antigenically Reactive Regions of the
TITLE OF INVENTION: Hepatitis A Virus Polypeptide
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick Stockton LLP
STREET: 3424 Peachtree Road, N.E.
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30326
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label= YK-1317
US-09-171-432A-41
Query Match 79.8%; Score 103; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 1 QRLKYAQEELSNEVLPPPRK 20
RESULT 11
US-09-171-432A-42
Sequence 42, Application US/09171432A
Publication No. US20030187184A1
GENERAL INFORMATION:
APPLICANT: Fields, Howard A.
APPLICANT: Khudiyakov, Yury E.
TITLE OF INVENTION: Antigenically Reactive Regions of the
TITLE OF INVENTION: Hepatitis A Virus Polypeptide
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick Stockton LLP
STREET: 3424 Peachtree Road, N.E.
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30326
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label= YK-1318
US-09-171-432A-42
Query Match 50.4%; Score 65; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 VLPPPRKMKGLF 25
DB 1 VLPPPRKMKGLF 12
RESULT 12
US-10-369-493-5713
Sequence 5713, Application US/10369493
Publication No. US2003033675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiaofeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 5713
LENGTH: 1193
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-5713
Query Match 40.3%; Score 52; DB 15; Length 1193;
Best Local Similarity 62.5%; Pred. No. .78;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVL 16
DB 340 QRLKYAQEELSNEVL 355
RESULT 13
US-10-369-493-4031

```
; Sequence 4031, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4031
; LENGTH: 1240
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1240)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-4031

Query Match      39.5%; Score 51; DB 15; Length 1240;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPP 18
   |||||:|||||
Db 253 QROSYASEDYENEVRGPP 270

RESULT 14
US-10-369-493-5013
; Sequence 5013, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5013
; LENGTH: 6642
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5013

Query Match      39.5%; Score 51; DB 15; Length 6642;
Best Local Similarity 45.5%; Pred. No. 7.1e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPPRKM 22
   :|||:|||||
Db 1353 RRVSPFAEELPKREVIDSRKK 1374

RESULT 15
US-10-369-493-3980
; Sequence 3980, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3980
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1168)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3980

Query Match      38.8%; Score 50; DB 15; Length 1168;
Best Local Similarity 47.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 YAQEELSNEVLPPPPRKM 21
   |||:|:|:|:|:|
Db 246 YAEEDSDELLPPRKM 262

Search completed: March 15, 2004, 13:53:28
Job time : 28.9706 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:50:08 ; Search time 12.5 Seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-09-171-432A-47
Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	341	2 S04137	genome polyprotein
2	129	100.0	852	1 GNNYHA	genome polyprotein
3	129	100.0	1358	2 A03905	genome polyprotein
4	129	100.0	2227	1 GNNYHM	genome polyprotein
5	129	100.0	2227	1 GNNYHR	genome polyprotein
6	124	96.1	2227	1 GNNYMK	genome polyprotein
7	123	95.3	2227	1 GNNYHB	genome polyprotein
8	119	92.2	2230	1 GNNYSA	genome polyprotein
9	117	90.7	839	1 GNNYS2	genome polyprotein
10	56.5	43.8	346	2 S74448	regulatory protein
11	52	40.3	859	2 T43701	DNA-directed RNA p
12	52	40.3	1193	2 E88445	protein C36E6.4 (i
13	51	39.5	1119	2 T50995	related to cytoske
14	51	39.5	6642	2 T29757	protein UNC-89 - C
15	50.5	39.1	443	2 E82046	proteinase HelvU,
16	49.5	38.4	159	2 C72210	conserved hypothet
17	49.5	38.4	554	2 A56730	carl protein - Pod
18	49	38.0	55	2 PQ0433	genome polyprotein
19	49	38.0	56	2 PQ0434	genome polyprotein
20	49	38.0	56	2 PQ0428	genome polyprotein
21	49	38.0	56	2 PQ0427	genome polyprotein
22	49	38.0	56	2 PQ0429	genome polyprotein
23	49	38.0	56	2 PQ0432	genome polyprotein
24	49	38.0	56	2 PQ0430	genome polyprotein
25	49	38.0	442	2 A71969	probable histidine
26	49	38.0	465	2 H86482	protein F5J5.11 (i
27	49	38.0	1174	2 S28976	DNA-directed RNA p
28	48.5	37.5	2278	1 S56274	PAB1 protein - yea
29	48	37.2	523	2 D85538	hypothetical prote

30 48 37.2 523 2 H90687
31 48 37.2 2 A27826
32 48 37.2 1191 2 S65068
33 48 37.2 1210 2 S35348
34 47.5 36.8 285 2 T15133
35 47.5 36.8 1626 2 T09271
36 47 36.4 223 2 PQ0514
37 47 36.4 254 2 F75575
38 47 36.4 578 2 S03299
39 47 36.4 592 1 LLBY
40 46.5 36.0 361 2 G82530
41 46.5 36.0 486 2 T10089
42 46.5 36.0 634 2 E86293
43 46 35.7 163 2 PC4186
44 46 35.7 262 2 A70428
45 46 35.7 338 2 I56893

ALIGNMENTS

RESULT 1

S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C/Species: human hepatitis A virus
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C/Accession: S04137
R/Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res 17, 3594, 1989
A/Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
A/Reference number: S04137; MUID:89263805; PMID:2542903
A/Accession: S04137
A/Molecule type: mRNA
A/Residues: 1-341 <AND>
A/Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576
C/Genetics:
A/Genes: VP1
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: Coat protein; polyprotein
P:2-340/Product: coat protein 1D (VP1) #status predicted <MAT>

Query Match 100.0%; Score 129; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. NO. 9.5e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 314 QRLKYAQEELSNEVLPPPRKMKGLF 338

RESULT 2

GNNYHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C/Species: human hepatitis A virus
A/Note: host Homo sapiens (man)
C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C/Accession: A03904
R/Linmeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.
J. Virol. 54, 247-255, 1985
A/Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A/Reference number: A03904; MUID:85185648; PMID:2985793
A/Accession: A03904
A/Molecule type: genomic RNA
A/Residues: 1-852 <LIN>
A/Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein, core protein, polyprotein
F:1-245/Product: coat protein 1A #status predicted <CIA>
F:246-491/Product: coat protein 1B #status predicted <CIB>
F:492-836/Product: coat protein 1C #status predicted <CIC>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

RESULT 4

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein

B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A>Note: host Homo sapiens (man)

C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R; Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with di

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: Genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus Genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2221/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 129; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 8,1e-11;

QY 1 QRLKYAQLSNEVLPPPRKMGGLF 25
| | | | |
Db 810 QRLKYAQLSNEVLPPPRKMGGLF 834

RESULT 6
GNMYM genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N; Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Notes: host Homo sapiens (man)
C>Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R; Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstone, S.M.; Purcell,
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison w/
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A>Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:G329594; PIDN:AAA45471.1; PID:G329595
A:Note: submitted to Genbank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidylyltra
F; 1-245/Product: coat protein 1A #status predicted <PIA>
F; 246-491/Product: coat protein 1B #status predicted <P1B>
F; 492-836/Product: coat protein 1C #status predicted <P1C>
F; 837-980/Product: core protein 2A #status predicted <P2A>
F; 981-1076/Product: core protein 2B #status predicted <P2B>
F; 1077-1422/Product: core protein 2C #status predicted <P2C>

F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 96.1%; Score 124; DB 1; Length 2227;
Best Local Similarity 96.0%; Pred. No. 4.6e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 7
GNVYB
genome polyprotein - human hepatitis A virus (strain MB8)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
Vg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J03003
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wisse, T.; Kiehn, R.; Wimmer, E.; Deinhardt,
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A:Reference number: J03003; MUID:68045071; PMID:2823500
A:Accession: J03003
A:Molecule type: Genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hyd
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPG>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 95.3%; Score 123; DB 1; Length 2227;
Best Local Similarity 96.0%; Pred. No. 6.6e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8
GNVYA
genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
submitted to JIPID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222597; PIDN:BAA00766.1; PID:g222598
R:Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an
A:Reference number: JQ1080; MUID:91311420; PMID:1649501
A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BALJ>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BALJ2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; core protein; polyprotein
F:1-27/Product: coat protein 1A #status predicted <C1A>
F:28-245/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: coat protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: protein 3A #status predicted <P3A>
F:1499-1521/Product: protein 3B #status predicted <P3B>
F:1522-1741/Product: protein 3C #status predicted <P3C>
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 92.2%; Score 119; DB 1; Length 2230;
Best Local Similarity 88.0%; Pred. No. 2.7e-09;
Matches 22; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 814 QRLKYAQEELSNEVLPPPRKMKGLF 838

RESULT 9
GNVY52
genome polyprotein - simian hepatitis A virus (strain CY-145) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat protei
C:Species: simian hepatitis A virus
A:Note: host Macaca fascicularis (cynomolgus macaque)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
C:Accession: JQ1180
R:Nainan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.
J. Gen. Virol. 72, 1685-1689, 1991
A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus ma
A:Reference number: JQ1180; MUID:91311421; PMID:1649902
A:Accession: JQ1180
A:Molecule type: genomic RNA
A:Residues: 1-839 <NAI>
A:Cross-references: GB:M59286; NID:g329599; PIDN:AAA5473.1; PID:g555083
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-23/Product: coat protein 1A #status predicted <VP0>
F:24-245/Product: coat protein 1B #status predicted <VP3>
F:246-491/Product: coat protein 1C #status predicted <VP1>
F:492-839/Product: core protein 2A (fragment) #status predicted <P2P>
F:839-1091/Product: core protein 2B (fragment) #status predicted <P2B>
F:1092-1426/Product: core protein 2C (fragment) #status predicted <P2C>
F:1427-1498/Product: protein 3A (fragment) #status predicted <P3A>
F:1499-1521/Product: protein 3B (fragment) #status predicted <P3B>
F:1522-1741/Product: protein 3C (fragment) #status predicted <P3C>
F:1742-2230/Product: protein 3D (fragment) #status predicted <P3D>

Query Match 90.7%; Score 117; DB 1; Length 839;
Best Local Similarity 84.0%; Pred. No. 1.8e-09;
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 809 QRLKYAQEELSNEVLPPPRKMKGLF 833

RESULT 10

S74448
regulatory protein pcrR - *Synechocystis* sp. (strain PCC 6803)
N:Alternate names: protein all1408
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
C:Accession: S74448
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3: 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74448
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-346 <KAN>
A:Cross-references: EMBL:D90899; GB:AB001339; NID:G1651650; PIDN:BAA16600.1; PID:G165167
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: pcrR
C:Superfamily: *Synechocystis* regulatory protein pcrR
C:Keywords: transcription regulation
Query Match 43.8%; Score 56.5; DB 2; Length 346;
Best Local Similarity 40.0%; Pred. No. 0.96;
Matches 14; Conservative 4; Mismatches 4; Indels 13; Gaps 1;
QY 1 QRLKYAQEELSNEVLPPP-----RQMK 22
DB 241 ERVKYAEILVKQLDPPSLAQLSRQVSLNRLK 275
RESULT 11
T43701
DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain [imported] - *Caenorhabditis el*
C:Species: *Caenorhabditis elegans*
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
C:Accession: T43701
R:Sidow, A.; Thomas, W.K.
Curr. Biol. 4: 596-603, 1994
A:Title: A molecular evolutionary framework for eukaryotic model organisms.
A:Reference number: Z22636; MUID:95041334; PMID:7953533
A:Accession: T43701
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-859 <SID>
A:Cross-references: EMBL:U10333; NID:G520506; PIDN:AAA50224.1; PID:G520507
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
C:Keywords: nucleotidyltransferase
Query Match 40.3%; Score 52; DB 2; Length 859;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPP 16
DB 141 QRKYAREILQKELLP 156
RESULT 12
E88445
protein C26E6.4 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: E88445
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gac/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E88445

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1193 <STO>
A:Cross-references: GB:chr_III; PIDN:AAA21158.1; PID:G532805; GSPDB:GN00021; CESP:C36E
C:Genetics:
A:Gene: C26E6.4
A:Map position: 3
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
Query Match 40.3%; Score 52; DB 2; Length 1193;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPP 16
DB 340 QRKYAREILQKELLP 355
RESULT 13
T50995
related to cytoskeleton assembly control protein SLA1 [imported] - *Neurospora crassa*
N:Alternate names: protein B7F18.140
C:Species: *Neurospora crassa*
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50995
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatur
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T50995
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1119 <SCH>
A:Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.140
A:Experimental source: BAC clone B7F18; strain OR74A
C:Genetics:
A:Gene: NCSP:B7F18.140
A:Map position: 6
A:Introns: 66/3; 123/2; 495/1
Query Match 39.5%; Score 51; DB 2; Length 1119;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPP 18
DB 165 QRQVASEDYENEVSPP 182
RESULT 14
T29757
protein UNC-89 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999
C:Accession: T29757
R:Du, Z.; Le, T.T.; Wilson, R.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of *C. elegans* cosmid C09D1.
A:Reference number: Z20679
A:Accession: T29757
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6642 <DUZ>
A:Cross-references: EMBL:AF003131; PIDN:AAE54132.1; GSPDB:GN00019; CESP:unc-89
A:Experimental source: strain Bristol N2; clone C09D1
C:Genetics:
A:Gene: CESP:unc-89
A:Map position: 1
A:Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1;
/3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1
Query Match 39.5%; Score 51; DB 2; Length 6642;
Best Local Similarity 45.5%; Pred. No. 1.9e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ORLKVAQEELSNEVLPPPRKMK 22
: : : : :
Db 1353 RVSFAEELPKVIDSDRKKK 1374

RESULT 15
E82046
Proteinase HslU, ATPase subunit HslU VC2674 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: E82046
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: E82046
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <HEI>
A:Cross-references: GB:AE004333; GB:AE003852; NID:G9657266; PIDN:AAF95815.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2674
A:Map position: 1
C:Superfamily: heat shock protein hslU; FtsH/SEC18/CDC48-type ATP-binding domain homolog

Query Match 39.1%; Score 50.5; DB 2; Length 443;
Best Local Similarity 39.3%; Pred. No. 10;
Matches 11; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 1 ORLKVAQEELSNE----VLPPPRKMKG 23
: : : : :
Db 117 EXVKFAEELAEERVLDALEPPPRDAWG 144

Search completed: March 15, 2004, 14:07:01
Job time : 13.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: March 15, 2004, 13:26:12 ; Search time 7.05882 Seconds
(without alignments)
184.415 Million cell updates/sec
Title: US-09-171-432A-47
Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_42.*

P17504 influenza b
P22032 influenza b
Q9zkt2 helicobacte
P44792 haemophilus
Q14849 homo sapien
Q17127 blaberus di
P38430 arabidopsis
O44757 caenorhabdi
P17896 drosophila
P11531 mus musculu
O97532 canis famil
P11532 homo sapien

ALIGNMENTS

RESULT 1
POLG_HPAV1 STANDARD; PRT; 341 AA.
AC P13672; 585 1 HEMA_INBE
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP3 (1C); Coat protein VP1 (1D); Core protein P2A] (Fragment).
OS Hepatitis A virus (strain LCDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
hepatitis A virus (HAV)."
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
EMBL; X14666; CAA32794.1; -;
PIR; S04137; S04137.
InterPro; IPR008975; Viral_cap_coat.
PolyProtein; Coat protein; Core protein.
KW NON TER 1 1
FT CHAIN <1 1 COAT PROTEIN VP3.
FT CHAIN 2 340 COAT PROTEIN VP1.
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;
Query Match 100.0%; Score 129; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 6.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 314 QRLKYAQEELSNEVLPPPRKMGGLF 338
RESULT 2
POLG_HPAVC STANDARD; PRT; 852 AA.
ID POLG_HPAVC

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	341	1 POLG_HPAV1	P13672 hepatitis a
2	129	100.0	852	1 POLG_HPAVC	P06442 hepatitis a
3	129	100.0	2226	1 POLG_HPAV2	P26580 hepatitis a
4	129	100.0	2226	1 POLG_HPAV4	P26581 hepatitis a
5	129	100.0	2226	1 POLG_HPAV8	P26582 hepatitis a
6	129	100.0	2227	1 POLG_HPAVH	P08617 hepatitis a
7	129	100.0	2227	1 POLG_HPAVL	P06441 hepatitis a
8	123	95.3	2227	1 POLG_HPAVM	P13901 hepatitis a
9	119	92.2	2230	1 POLG_HPAVS	P14553 simian hepa
10	117	90.7	839	1 POLG_HPAVT	P31788 simian hepa
11	103	79.8	808	1 POLG_HPAVG	Q02381 hepatitis a
12	53	41.1	636	1 NAF1_HUMAN	Q15025 homo sapien
13	52	40.3	1193	1 RP22_CABEL	Q10578 caenorhabdi
14	51	39.5	6632	1 UN89_CABEL	O01761 caenorhabdi
15	50.5	39.1	443	1 HSLU_VIBCH	Q9knq7 vibrio chol
16	50.5	39.1	443	1 HSLU_VIEVU	O8dcp4 vibrio vuln
17	49.5	38.4	554	1 PEX2_PODAN	P51021 podospora a
18	49	38.0	1174	1 RP22_HUMAN	P30876 homo sapien
19	48.5	37.6	2278	1 FAB1_YEAST	P34756 saccharomyc
20	48	37.2	1176	1 RP22_DROME	P08266 drosophila
21	48	37.2	1191	1 RP22_LYCES	Q42877 lycopersico
22	48	37.2	1210	1 RP22_SCHPO	Q02061 schizosach
23	47	36.4	578	1 HEMA_INBME	P09765 influenza b
24	47	36.4	592	1 ABP1_YEAST	P15891 saccharomyc
25	46.5	36.0	399	1 BM8A_MOUSE	P34821 mus musculu
26	46	35.7	338	1 TAP4_HUMAN	Q10664 homo sapien
27	46	35.7	480	1 I1SB_CUCMA	P13744 cucurbita m
28	46	35.7	574	1 HEMA_INBMD	P03461 influenza b
29	46	35.7	576	1 HEMA_INEUS	P09766 influenza b
30	46	35.7	578	1 HEMA_INEVI	P09767 influenza b
31	46	35.7	583	1 HEMA_INEIN	P10757 influenza b
32	46	35.7	583	1 HEMA_INBOR	P03464 influenza b
33	46	35.7	583	1 HEMA_INBSI	P03463 influenza b

AC P06442; Q83741; Q83742;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment)
OS Hepatitis A virus (strain CR326).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
CX NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95185648; PubMed=2985793;
RA Linemeyer D.L., Wenke J.G., Martin-Gallardo A., Hughes J.V.,
RA Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
RL J. Virol. 54:247-255(1985).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M10033; AAA45470.1; -;
DR PIR; A03904; GNNYHA.
DR InterPro; IPR008975; Viral_cap_coat.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
FT CHAIN 837 >852 CORE PROTEIN P2A.
FT NON_TER 852 852
SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;

Query Match 100.0%; Score 129; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 3
POLG HPAV2
ID POLG HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
CX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
RA Cromean T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic

RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59810; AAA45468.1; -;
DR MEROPS; C03.005; -;
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D634E2BF CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 4
POLG HPAV4
ID POLG HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
CX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
RA Cromean T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic

RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59809; AAA45469.1; -
CC MEROPS; C03.005; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR006055; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC PolyProtein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
CC FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9BF75 CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
|||
DB 810 QRLKYAQEELSNEVLPPPRKMGFLF 834
|||

RESULT 5
POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91182758; PubMed=1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59808; AAA45467.1; -
CC PDB; 1QA7; 15-MAY-00.
CC MEROPS; C03.005; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR006055; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC PolyProtein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
CC FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
|||
DB 810 QRLKYAQEELSNEVLPPPRKMGFLF 834
|||

RESULT 6
POLG_HPAV8 STANDARD; PRT; 2227 AA.
AC P08617; P06443; Q81082;
DT 01-AUG-1998 (Rel. 08, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12098;
OX

RN [1] SEQUENCE FROM N.A.
 RP STRAIN=Wild type;
 RX MEDLINE=87061253; PubMed=3023706;
 RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
 RA Baroudy B.M.;
 RT "Complete nucleotide sequence of wild-type hepatitis A virus;
 RT comparison with different strains of hepatitis A virus and other
 RT picornaviruses";
 RL J. Virol. 61:50-59(1987).
 RN [2] SEQUENCE FROM N.A.
 RP STRAIN=Attenuated;
 RX MEDLINE=87175701; PubMed=3031686;
 RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
 RA Purcell R.H.;
 RT "Complete nucleotide sequence of an attenuated hepatitis A virus;
 RT comparison with wild-type virus";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
 RN [3] SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
 RP MEDLINE=8516289; PubMed=2984684;
 RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
 RA Purcell R.H., Feinstone S.M.;
 RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
 RT proteins and RNA polymerase";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}[N]
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
 CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -!- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
 CC SHOWN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M14114; AAA45475.1; -;
 CC EMBL; M14707; AAA45465.1; -;
 CC EMBL; M14707; AAA45466.1; ALT_INT.
 CC EMBL; M16632; AAA45471.1; -;
 CC PIR; A03905; A03905.
 CC PIR; A25981; GNNYHM.
 CC PIR; A94149; GNNYMK.
 CC PDB; 1HAV; 23-DEC-96.
 CC MEROPS; C03.005; -;
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_ser_trypsin.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol_1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUSNS.
 CC PolyProtein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; TRANSFERASE; 3D-STRUCTURE.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN CORE PROTEIN P2A.

FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 FT VARIANT 77 77
 FT VARIANT 764 764
 FT VARIANT 821 821
 FT VARIANT 1052 1052
 FT VARIANT 1062 1062
 FT VARIANT 1118 1118
 FT VARIANT 1151 1151
 FT VARIANT 1163 1163
 FT VARIANT 1277 1277
 FT VARIANT 1500 1500
 FT VARIANT 1805 1805
 FT VARIANT 1930 1930
 SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 5.2e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQBELSNEVLPPPRKMKGLF 25
 DB 810 QRLKYAQBELSNEVLPPPRKMKGLF 834
 RESULT 7
 POLG HPAVL STANDARD; PRT; 2227 AA.
 ID POLG HPAVL STANDARD; PRT; 2227 AA.
 AC P06411;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain LA).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12099;
 RN [1]_TaxID=12099;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85190549; PubMed=2986127;
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
 RA Merryweather J., van Nest G., Dina D.;
 RT "Primary structure and gene organization of human hepatitis A virus";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}[N]
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K02990; AAA45472.1; -;
 CC FIR; A03903; GNNYHR.
 CC MEROPS; C03.005; -;
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_ser_trypsin.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol_1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUSNS.
 CC PolyProtein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; TRANSFERASE; 3D-STRUCTURE.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN CORE PROTEIN P2A.

```
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_P3D.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8
POLG HPVAVM STANDARD; PRT: 2227 AA.
AC P13901; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MBB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wisse T., Kiehn R., Wimmer E.,
RA Deinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
RT virus (isolate MBB).";
RL Virus Res. 8:153-171(1987).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M20273; AAA45474.1; -.
CC MEROPS; C03.005; -.

DR InterPro: IPR004004; Calici_pol_hel.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR00605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_P3vir.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 95.3%; Score 123; DB 1; Length 2227;
Best Local Similarity 96.0%; Pred. No. 4.2e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 9
POLG HPVAVS STANDARD; PRT: 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tearav S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjapardize A.G., Tearav S.A.,
RA Sverdlov E.D., Chazhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00924; BAA00766.1; -;
CC DR EMBL; X15461; CRA33490.1; -;
CC DR PIR; A30470; GNNYS2.
CC
CC MEROPS; C03.005; -;
CC DR InterPro; IPR004004; Calici_pol_hel.
CC DR InterPro; IPR009003; Cys_Ser_trypsin.
CC DR InterPro; IPR006005; RNA_helicase.
CC DR InterPro; IPR007095; RNA_pol_DS_PS.
CC DR InterPro; IPR001205; RNA_pol_P3D.
CC DR InterPro; IPR007094; RNA_pol_PSVir.
CC DR InterPro; IPR008975; Viral_cap_coat.
CC DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam; PF00910; RNA_helicase; 1.
CC DR PRINIS; PR00918; CALICIVIRUSNS.
CC
CC KW Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
CC FT CHAIN 1 27
CC FT CHAIN 28 249
CC FT CHAIN 250 495
CC FT CHAIN 496 795
CC FT CHAIN 796 984
CC FT CHAIN 985 1091
CC FT CHAIN 1092 1426
CC FT CHAIN 1427 1498
CC FT CHAIN 1499 1521
CC FT CHAIN 1522 1741
CC FT CHAIN 1742 2230
CC FT CHAIN 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
CC SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 92.2%; Score 119; DB 1; Length 2230;
Best Local Similarity 88.0%; Pred. No. 1.7e-09;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 814 QRLKYAEELSNEILPPPRKVKGLF 838

RESULT 10
POLG HPVAVT
ID POLG HPVAVT STANDARD; PRT; 839 AA.
AC P31788;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein (Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A) (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
RT cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M6695; AAA45477.1; -;
CC DR InterPro; IPR008975; Viral_cap_coat.
CC KW Polyprotein; Coat protein; Core protein.
CC FT NON_TER 1 1
CC FT CHAIN <1 2
CC FT CHAIN 3 223
CC FT CHAIN 224 470
CC FT CHAIN 471 770
CC FT CHAIN 771 >808
CC FT NON_TER 808 808
CC SQ SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;

Query Match 79.8%; Score 103; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59286; AAA45473.1; -;
CC DR PIR; JQ1180; GNNYS2.
CC
CC KW Polyprotein; Coat protein; Core protein.
CC FT CHAIN 1 23
CC FT CHAIN 24 245
CC FT CHAIN 246 491
CC FT CHAIN 492 ?
CC FT CHAIN ? >839
CC FT NON_TER 839 839
CC SQ SEQUENCE 839 AA; 93825 MW; 2CACCA4BD1E192DBC CRC64;

Query Match 90.7%; Score 117; DB 1; Length 839;
Best Local Similarity 84.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 809 QRFKYARELSNEILPPPRKLGFLF 833

RESULT 11
POLG HPVAVT
ID POLG HPVAVT STANDARD; PRT; 808 AA.
AC Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein (Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A) (Fragment).
OS Hepatitis A virus (strain GA76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;
RT "Characterization of a Genetic variant of human hepatitis A virus.";
RL J. Med. Virol. 36:118-124(1992).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M6695; AAA45477.1; -;
CC DR InterPro; IPR008975; Viral_cap_coat.
CC KW Polyprotein; Coat protein; Core protein.
CC FT NON_TER 1 1
CC FT CHAIN <1 2
CC FT CHAIN 3 223
CC FT CHAIN 224 470
CC FT CHAIN 471 770
CC FT CHAIN 771 >808
CC FT NON_TER 808 808
CC SQ SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;

Query Match 79.8%; Score 103; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
 DB 789 QRLKYAQEELSNEVLPPPRK 808

RESULT 12
 ID NAF1 HUMAN STANDARD; PRT: 636 AA.
 AC Q15025; Q76008; Q96EL9; Q99833; Q9H1J3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Nef-associated factor 1 (Naf1) (HIV-1 Nef interacting protein)
 DE (Virion-associated nuclear shuttling protein) (VAN) (hVAN) (TNFAIP3
 DE interacting protein 1).
 GN TNIP1 OR NAF1 OR KIAA0113.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Peripheral blood;
 RX MEDLINE=99120485; PubMed=9923610;
 RA Fukushima M., Dixon J., Kimura T., Tsurutani N., Dixon M.J.,
 RA Yamamoto N.;
 RT "Identification and cloning of a novel cellular protein Naf1, Nef-
 RT associated factor 1, that increases cell surface CD4 expression";
 RL PNAS Lett. 442:83-88(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Leukocyte;
 RX MEDLINE=20541981; PubMed=11090181;
 RA Gupta K., Ott D., Hope T.J., Siliciano R.F., Boeke J.D.;
 RT "A human nuclear shuttling protein that interacts with human
 RT immunodeficiency virus type 1 matrix is packaged into virions";
 RL J. Virol. 74:11811-11824(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 136-636 FROM N.A. (ISOFORM 2).
 RC TISSUE=Craniofacial;
 RX MEDLINE=96276047; PubMed=8691136;
 RA Loftus S.K., Dixon J., Koprivnikar K., Dixon M.J., Wasmuth J.J.;
 RT "Transcriptional map of the Treacher Collins candidate gene region";
 RL Genome Res. 6:26-34(1996).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Chara O., Nagase T., Kikuno R., Nomura N.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

RN RP SEQUENCE OF 341-636 FROM N.A. (ISOFORM 1).
 RC TISSUE=Bone marrow;
 RX MEDLINE=95309325; PubMed=7785527;
 RA Nagase T., Miyajima N., Tanaka A., Suzuki T., Seki N., Sato S.,
 RA Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. III.
 RT The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 2:37-43(1995).
 RN [7]
 RP SEQUENCE OF 94-412 FROM N.A.
 RA Fukushi M., Kimura T., Yamamoto N.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Interacts with zinc finger protein A20/TNFAIP3 and
 CC inhibits TNF-induced NF-kappa-B-dependent gene expression by
 CC interfering with an RIP- or TRAF2-mediated transactivation signal
 CC (By similarity). Increases cell surface CD4(T4) antigen
 CC expression. Interacts with HIV-1 matrix protein and is packaged
 CC into virions and overexpression can inhibit viral replication. May
 CC regulate matrix nuclear localization, both nuclear import of PIC
 CC (preintegration complex) and export of GAG polypeptide and viral
 CC genomic RNA during virion production.
 CC -1- SUBUNIT: Interacts with TNFAIP3 (By similarity). Interacts with
 CC HIV-1 matrix protein.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Shuttles between the nucleus
 CC and cytoplasm in a CRM1-dependent manner.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Alpha;
 CC IsoId=Q15025-1; Sequence=Displayed;
 CC Name=2; Synonyms=Beta;
 CC IsoId=Q15025-2; Sequence=VSP_003913;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Strongly expressed in peripheral
 CC blood lymphocytes, spleen and skeletal muscle, and is weakly
 CC expressed in the brain.
 CC -1- CAUTION: Ref.7 sequence differs from that shown due to frameshifts
 CC in positions 152 and 154.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AJ011895; CAA09855.1; -
 CC EMBL; AJ011896; CAA09856.1; -
 CC EMBL; AY012155; AAG42154.1; -
 CC EMBL; BC012133; AAH12133.1; -
 CC EMBL; BC014008; AAH14008.1; -
 CC EMBL; U39403; AAC99999.1; -
 CC EMBL; D30755; BAA06416.2; -
 CC EMBL; U83844; AAB41438.1; ALT_FRAME.
 CC Genew; HGNC:16903; TNIP1.
 CC MIM; 607714; -
 CC GO; GO:0005622; C:intracellular; TAS.
 CC GO; GO:0005515; F:protein binding; TAS.
 CC GO; GO:0009101; F:glycoprotein biosynthesis; IDA.
 CC GO; GO:0045071; P:negative regulation of viral genome replica...; TAS.
 CC Coiled coil; Nuclear protein; Alternative splicing.
 CC DOMAIN 20 73 COILED COIL (POTENTIAL).
 CC DOMAIN 196 258 COILED COIL (POTENTIAL).
 CC DOMAIN 294 535 COILED COIL (POTENTIAL).
 CC DOMAIN 94 412 INTERACTS WITH NEF.
 CC DOMAIN 524 530 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC DOMAIN 539 636 PRO-RICH.
 CC VARSPLIC 627 636 SPKNDREGPO -> PADLRPRN (in isoform 2).
 CC /FTID=VSP_003913.
 CC G -> D (IN REF. 3; AAH12133).
 CC CONFLICT 148 148 A -> P (IN REF. 2).
 CC CONFLICT 299 299

SQ SEQUENCE 636 AA; 71864 MW; D81B96BEAD50D871 CRC64;
Query Match 41.1%; Score 53; DB 1; Length 636;
Best Local Similarity 50.0%; Pred. No. 3.7;
Matches 13; Conservative 2; Mismatches 9; Indels 2; Gaps 1;
QY 2 RLKYAQEEL--SNEVLPPPRKWKGLF 25
|||:|||||:|||||:|||||
Db 58 RLKQKABELVKDNELLPPPSLGSF 83
RESULT 13
RPB2 CAEEL STANDARD; PRT; 1193 AA.
AC Q10578;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DNA-directed RNA polymerase II second largest subunit (EC 2.7.7.6)
DE (RNA polymerase I subunit 2).
GN C26E6.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulton L.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 200-1058 FROM N.A.
RX MEDLINE=95041334; PubMed=7953533;
RA Sidow A., Thomas W.K.;
RT "A molecular evolutionary framework for eukaryotic model organisms.";
RL Curr. Biol. 4:596-603(1994).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
of DNA into RNA using the four ribonucleoside triphosphates as
substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC -!- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14
different polypeptides. This subunit is the second largest
component of RNA polymerase II.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
found in eukaryotic nuclei: polymerase I for the ribosomal RNA
precursor, polymerase II for the mRNA precursor, and polymerase
III for 5S and tRNA genes.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U13875; AAA21158.1; -
CC EMBL; U10333; AAA50224.1; -
CC PIR; E88445; E88445.
CC PIR; T43701; T43701.
CC WormPep; C26E6.4; CE01162.
CC InterPro; IPR007121; RNA_pol_B.
CC InterPro; IPR007644; RNA_pol_Rpb2_1.
CC InterPro; IPR007642; RNA_pol_Rpb2_2.
CC InterPro; IPR007645; RNA_pol_Rpb2_3.
CC InterPro; IPR007646; RNA_pol_Rpb2_4.
CC InterPro; IPR007647; RNA_pol_Rpb2_5.
CC InterPro; IPR007120; RNA_pol_Rpb2_6.
CC InterPro; IPR007641; RNA_pol_Rpb2_7.
CC Pfam; PF04563; RNA_pol_Rpb2_1; -
DR Pfam; PF04561; RNA_pol_Rpb2_2; 1.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT ZN FING 1125 1146 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1193 AA; 134904 MW; B8A85E74E9CCTEBE CRC64;
Query Match 40.3%; Score 52; DB 1; Length 1193;
Best Local Similarity 62.5%; Pred. No. 10; Mismatches 3; Indels 0; Gaps 0;
Matches 10; Conservative 3;
QY 1 ORLYAQEELSNEVLPP 16
|||:|||||:|||||
Db 340 QRIKYAREILQKELLP 355
RESULT 14
UN89 CAEEL STANDARD; PRT; 6632 AA.
AC U01761; Q17362;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
DE UNC-89 OR C09D1.1.
GN Caenorhabditis elegans.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Bristol N2;
RX MEDLINE=96180278; PubMed=8603916;
RA Benlan G.M., Tinley T.L., Tang X., Borodovsky M.;
RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line
transduction domains.";
RL J. Cell Biol. 132:835-848(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du Z., Le T.T., Wilson R.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Structural component of the muscle M-line. Myofibrilament
lattice assembly begins with positional cues laid down in the
basement membrane and muscle cell membrane. UNC-89 responds to
these signals, localizes, and then participates in assembling an
M-line.
CC -!- TISSUE SPECIFICITY: Localizes to the middle of A-bands.
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -!- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 5 RCD domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

EMBL; U33058; AAB00542.1; -.
DR EMBL; AF003131; AAB54132.2; -.
DR PDB; IFHO; 2O-DEC-00.
DR WormPep; C09D1.1; CE30426.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003066; Ig_VHC.
DR InterPro; IPR001849; PH.
DR InterPro; IPR007850; RCSD.
DR InterPro; IPR000219; RhGEF.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 47.
DR Pfam; PF00169; Ph; 1.
DR Pfam; PF05177; RCSD; 5.
DR Pfam; PF00621; RhGEF; 1.
DR Pfam; PF00018; SH3; 1.
DR SMART; SMC0408; IGc2; 23.
DR SMART; SMO0325; RhGEF; 1.
DR SMART; SMO0326; SH3; 1.
DR PROSITE; PS0010; DH 2; 1.
DR PROSITE; PS0835; IG-LIKE; 49.
DR PROSITE; PS00003; RH_DOMAIN; 1.
DR PROSITE; PS00002; SH3; 1.
KW Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
3D-structure.

FT DOMAIN 63 127
FT DOMAIN 152 330
FT DOMAIN 342 498
PH.
FT DOMAIN 547 633
IG-LIKE C2-TYPE 1.
FT DOMAIN 648 736
IG-LIKE C2-TYPE 2.
FT DOMAIN 748 838
IG-LIKE C2-TYPE 3.
FT DOMAIN 946 1033
IG-LIKE C2-TYPE 4.
FT DOMAIN 1044 1132
IG-LIKE C2-TYPE 5.
FT DOMAIN 1140 1227
IG-LIKE C2-TYPE 6.
THR-RICH.
FT DOMAIN 1272 1315
RCSD 1.
FT DOMAIN 1375 1475
RCSD 2.
FT DOMAIN 1479 1585
RCSD 3.
FT DOMAIN 1597 1695
RCSD 4.
FT DOMAIN 1700 1799
RCSD 5.
FT DOMAIN 1800 1860
RCSD 6.
FT DOMAIN 1882 2067
IG-LIKE C2-TYPE 7.
FT DOMAIN 2071 2163
IG-LIKE C2-TYPE 8.
FT DOMAIN 2171 2261
IG-LIKE C2-TYPE 9.
FT DOMAIN 2269 2359
IG-LIKE C2-TYPE 10.
FT DOMAIN 2367 2455
IG-LIKE C2-TYPE 11.
FT DOMAIN 2463 2564
IG-LIKE C2-TYPE 12.
FT DOMAIN 2563 2651
IG-LIKE C2-TYPE 13.
FT DOMAIN 2657 2746
IG-LIKE C2-TYPE 14.
FT DOMAIN 2754 2858
IG-LIKE C2-TYPE 15.
FT DOMAIN 2887 2980
IG-LIKE C2-TYPE 16.
FT DOMAIN 2994 3081
IG-LIKE C2-TYPE 17.
FT DOMAIN 3087 3183
IG-LIKE C2-TYPE 18.
FT DOMAIN 3189 3280
IG-LIKE C2-TYPE 19.
FT DOMAIN 3286 3376
IG-LIKE C2-TYPE 20.
FT DOMAIN 3384 3469
IG-LIKE C2-TYPE 21.
FT DOMAIN 3482 3572
IG-LIKE C2-TYPE 22.
FT DOMAIN 3580 3667
IG-LIKE C2-TYPE 23.
FT DOMAIN 3686 3777
IG-LIKE C2-TYPE 24.
FT DOMAIN 3817 3908
IG-LIKE C2-TYPE 25.
FT DOMAIN 3920 4009
IG-LIKE C2-TYPE 26.
FT DOMAIN 4018 4106
IG-LIKE C2-TYPE 27.
FT DOMAIN 4109 4201
IG-LIKE C2-TYPE 28.
FT DOMAIN 4212 4297
IG-LIKE C2-TYPE 29.
FT DOMAIN 4302 4387
IG-LIKE C2-TYPE 30.
FT DOMAIN 4400 4485
IG-LIKE C2-TYPE 31.
FT DOMAIN 4489 4580
IG-LIKE C2-TYPE 32.
FT DOMAIN 4588 4678
IG-LIKE C2-TYPE 33.
FT DOMAIN 4681 4771
IG-LIKE C2-TYPE 34.
FT DOMAIN 4873 4961
IG-LIKE C2-TYPE 35.
FT DOMAIN 4965 5057
IG-LIKE C2-TYPE 36.

[illegible]

RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Seilers P.,
 RA McDonald L., Uterback T., Fieschmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.,
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*."; Nature 406:477-483 (2000).
 RL Nature 406:477-483 (2000).
 CC -!- FUNCTION: Chaperone subunit of a proteasome-like degradation
 CC complex (By similarity).
 CC -!- SUBUNIT: A double ring-shaped homohexameric of hslv is capped on
 CC each side by a ring-shaped hslu homohexameric (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the clpX chaperone family. Hslu subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE004333; AAF95815.1; -.
 DR PIR; E82046; E82046.
 DR HSP; P32168; IDO2.
 DR TIGR; VC2674; -.
 DR HAMAP; MF 00249; -. 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003599; AAA_ATPase_central.
 DR InterPro; IPR004491; Hsp_Hslvu.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRPFAMS; TIGR00390; hslu; 1.
 KW Chaperone; ATP-binding; Complete proteome.
 FT NP BIND 57 64 ATP (POTENTIAL).
 SQ SEQUENCE 443 AA; 49900 MW; DAE13E82FAB6A38F CRC64;

Query Match 39.1%; Score 50.5; DB 1; Length 443;
 Best Local Similarity 39.3%; Pred. No. 5.9;
 Matches 11; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 1 QRLKYAQEELSNE-----VLPPEKVKKG 23
 Db 117 EKVKFRAEELAEERVLDAIPPRDANG 144

Search completed: March 15, 2004, 14:01:01
 Job time : 8.05882 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:49:12 ; Search time 36.4706 Seconds
(without alignments)
216.283 Million cell updates/sec

Title: US-09-171-432A-47
Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL 25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_tvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	94	12 Q9ENT9	Q9ent9 hepatitis a
2	129	100.0	94	12 Q9ENU3	Q9enu3 hepatitis a
3	129	100.0	94	12 Q9ENV9	Q9env9 hepatitis a
4	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
5	129	100.0	94	12 Q9ENU2	Q9enu2 hepatitis a
6	129	100.0	94	12 Q9ENT1	Q9ent1 hepatitis a
7	129	100.0	94	12 Q9ENU1	Q9enu1 hepatitis a
8	129	100.0	94	12 Q9ENU8	Q9enu8 hepatitis a
9	129	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
10	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
11	129	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
12	129	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
13	129	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
14	129	100.0	94	12 Q9ENV8	Q9env8 hepatitis a
15	129	100.0	94	12 Q9ENV6	Q9env6 hepatitis a
16	129	100.0	94	12 Q9ENW0	Q9enw0 hepatitis a

17	129	100.0	94	12 Q9ENU6	Q9enu6 hepatitis a
18	129	100.0	94	12 Q9ENU4	Q9enu4 hepatitis a
19	129	100.0	94	12 Q9ENW1	Q9enw1 hepatitis a
20	129	100.0	94	12 Q9ENT5	Q9ent5 hepatitis a
21	129	100.0	94	12 Q9ENU7	Q9enu7 hepatitis a
22	129	100.0	94	12 Q9ENV4	Q9env4 hepatitis a
23	129	100.0	115	12 Q9DWR4	Q9dwr4 hepatitis a
24	129	100.0	116	12 Q92941	Q92941 hepatitis a
25	129	100.0	116	12 Q9W7X7	Q9w7x7 hepatitis a
26	129	100.0	116	12 Q9W7S7	Q9w7s7 hepatitis a
27	129	100.0	116	12 Q71977	Q71977 hepatitis a
28	129	100.0	116	12 Q8B8K6	Q8b8k6 hepatitis a
29	129	100.0	116	12 Q8B8K5	Q8b8k5 hepatitis a
30	129	100.0	116	12 Q8B8K4	Q8b8k4 hepatitis a
31	129	100.0	116	12 Q8B8K3	Q8b8k3 hepatitis a
32	129	100.0	126	12 Q9BW19	Q9bw19 hepatitis a
33	129	100.0	132	12 Q8V4L9	Q8v4l9 hepatitis a
34	129	100.0	132	12 Q8V4L6	Q8v4l6 hepatitis a
35	129	100.0	132	12 Q8V4J9	Q8v4j9 hepatitis a
36	129	100.0	132	12 Q8V4J5	Q8v4j5 hepatitis a
37	129	100.0	132	12 Q8V4M5	Q8v4m5 hepatitis a
38	129	100.0	132	12 Q8V4M7	Q8v4m7 hepatitis a
39	129	100.0	132	12 Q8V4N2	Q8v4n2 hepatitis a
40	129	100.0	132	12 Q8V4M4	Q8v4m4 hepatitis a
41	129	100.0	132	12 Q8V4J7	Q8v4j7 hepatitis a
42	129	100.0	132	12 Q8V4L5	Q8v4l5 hepatitis a
43	129	100.0	132	12 Q8V4K3	Q8v4k3 hepatitis a
44	129	100.0	132	12 Q8V4L0	Q8v4l0 hepatitis a
45	129	100.0	132	12 Q8V4M2	Q8v4m2 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENT9 PRELIMINARY; PRT; 94 AA.
AC Q9ENT9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Polypeptide (fragment).
OS Hepatitis A virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 32;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038298; BAB11836.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10807 MW; F194CE91BB8C4AFD CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Db 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 2

Q9ENU3

ID Q9ENJ3 PRELIMINARY; PRT; 94 AA.
 AC Q9ENJ3; 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 27;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038294; BAB11832.1; -
 DR PIR; PQ0427; PQ0428;
 DR PIR; PQ0428; PQ0428;
 DR PIR; PQ0430; PQ0430;
 DR PIR; PQ0431; PQ0431;
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10892 MW; 9AF8BE91BB8C53CC CRC64;
 Query Match 100.0%; Score 129; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 3.8e-12; Indels 0; Gaps 0;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 3

ID Q9ENV9 PRELIMINARY; PRT; 94 AA.
 AC Q9ENV9; 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 03;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038278; BAB11816.1; -
 DR PIR; PQ0427; PQ0428;
 DR PIR; PQ0428; PQ0428;
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10862 MW; 9AF9EFD4B8C4BAC CRC64;
 Query Match 100.0%; Score 129; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 3.8e-12; Indels 0; Gaps 0;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 4

ID Q9ENV5 PRELIMINARY; PRT; 94 AA.
 AC Q9ENV5; 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 08;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038282; BAB11820.1; -
 DR PIR; PQ0427; PQ0427;
 DR PIR; PQ0428; PQ0428;
 DR PIR; PQ0430; PQ0430;
 DR PIR; PQ0431; PQ0431;
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 129; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 3.8e-12; Indels 0; Gaps 0;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 5

ID Q9ENU2 PRELIMINARY; PRT; 94 AA.
 AC Q9ENU2; 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 28;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038295; BAB11833.1; -
 DR PIR; PQ0427; PQ0427;
 DR PIR; PQ0428; PQ0428;
 DR PIR; PQ0430; PQ0430;
 DR PIR; PQ0431; PQ0431;
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 129; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 3.8e-12; Indels 0; Gaps 0;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

Query Match 100.0%; Score 129; DB 12; Length 94;

```
Query Match      100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Caps 0;
```

Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 10
Q9ENV5 Q9ENV5 PRELIMINARY; PRT; 94 AA.
ID Q9ENV5
AC Q9ENV5
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 25;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038292; BAB11830.1; -;
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 11
Q9ENV1 Q9ENV1 PRELIMINARY; PRT; 94 AA.
ID Q9ENV1
AC Q9ENV1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 18;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038286; BAB11824.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 12
Q9ENV9 Q9ENV9 PRELIMINARY; PRT; 94 AA.
ID Q9ENV9
AC Q9ENV9
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 21;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038288; BAB11826.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 13
Q9ENV7 Q9ENV7 PRELIMINARY; PRT; 94 AA.
ID Q9ENV7
AC Q9ENV7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 06;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038280; BAB11818.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.

```
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMGGLF 85
RESULT 14
Q9ENV8
ID Q9ENV8 PRELIMINARY; PRT; 94 AA.
AC Q9ENV8;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
CX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 04;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the ENBL/GenBank/DBJ databases.
DR ENBL; AB038279; BAB11817.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMGGLF 85
RESULT 15
Q9ENV6
ID Q9ENV6 PRELIMINARY; PRT; 94 AA.
AC Q9ENV6;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
CX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 07;
RA Ida S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the ENBL/GenBank/DBJ databases.
DR ENBL; AB038281; BAB11819.1; -.
DR PIR; PQ0427; PQ0427.
```

```
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMGGLF 85
```

Search completed: March 15, 2004, 14:05:22
JOB time : 36.4706 secs